Universiteit Antwerpen

Letteren en Wijsbegeerte

Departement Taalkunde

The pluralization of presentational *haber* in Caribbean Spanish.

A study in Cognitive Construction Grammar and Comparative Sociolinguistics

Proefschrift voorgelegd tot het behalen van de graad van Doctor in de Taalkunde aan de Universiteit Antwerpen te verdedigen door:

JEROEN CLAES

Prof. Dr. Frank Brisard
Prof. Dr. Nicole Delbecque

Antwerpen, 2014

For Sara

[L]inguistic theory can no more ignore the social behavior of speakers of a language than chemical theory can ignore the observed properties of elements. The penalties for ignoring data from the speech community are a growing sense of frustration, a proliferation of moot questions, and a conviction that linguistics is a game in which each theorist chooses the solution that fits his taste or intuition. I do not believe that we need at this point a new 'theory of language'; rather, we need a new way of doing linguistics that will yield decisive solutions – WILLIAM LABOV.[1]

The dialectology of the New World offers an attractive opportunity to study linguistic changes in progress. ... As we follow their antecedents backwards in time, we encounter the dialectology and language contacts of the Old World, where layers of intersecting influence accumulate over the centuries. The record is blurred and many times overlaid, but it is worth deciphering. Tracing history as it is being made is exhilarating, but it is always helpful to know where we came from – WILLIAM LABOV.[2]

[1] Labov (1972: 259).
[2] Labov (2010: 119).

The pluralization of presentational *haber* in Caribbean Spanish

A STUDY IN COGNITIVE CONSTRUCTION GRAMMAR AND COMPARATIVE SOCIOLINGUISTICS

DE PLURALISERING VAN PRESENTATIONEEL *HABER* IN HET CARIBISCH SPAANS. EEN STUDIE IN COGNITIEVE CONSTRUCTIEGRAMMATICA EN VERGELIJKENDE SOCIOLINGUÏSTIEK

Jeroen Claes

Advisors:
Prof. Dr. Frank Brisard
Prof. Dr. Nicole Delbecque

University of Antwerp
Faculty of Arts, Department of Linguistics
2014

This page intentionally left blank

Contents

List of tables

List of figures

Abbreviations and other conventions

< >	Construction schema
[ ]	Literal translation
Acc	Accusative case
AdvP	Adverbial Phrase
boldface	Profiled portions of event frames
LH01H22/LH33	The codes at the end of the examples identify the cases in my corpus: LH=Havana (SD=Santo Domingo, SJ=San Juan); 01= informant number 1; H=male informant (M=female); 2=55+ years of age (1=20-35 years of age); 2=university graduate (1=less than university). The code behind the backslash is the identifier of the example in the database.[1]
Nom	Nominative case
NP	Noun Phrase
Obj	Direct object of a one- or two-argument construction
Obj1	Indirect object of a three-argument construction
Obj2	Direct object of a three-argument construction
Plur	Pluralized presentational *haber*
PP	Prepositional Phrase
Sing	Singular presentational *haber*
Subj	Subject

[1] See Chapter 4.1 and Chapter 4.2 for discussion of the methods that were used in gathering the corpus.

Acknowledgements

Many people have contributed to the success of this project. For example, Professor Luis A. Ortiz-López (*Universidad de Puerto Rico, Recinto de Río Piedras*) and his research assistant Nadja Fuster have done a great job looking for participants in San Juan. Professor Sunny Cabrera-Salcedo, the Chair of the Graduate Linguistics Program at the *Universidad de Puerto Rico, Recinto de Río Piedras*, was kind enough to write a letter of recommendation that opened many doors. Her help in locating participants was also appreciated. In addition, I would like to thank Boris Carrasquillo, José-Alberto Santiago, and Naiska Guzmán for driving me around San Juan to interview friends, acquaintances, or family of theirs. The *Casas Culturales* of the Municipality of San Juan in the boroughs of Tras Talleres and Chícharo were also a great help, as were the homeless shelter/rehabilitation center *Fondita de Jesús* and the elderly residence *Sagrado Corazón*. Finally, I am indebted with Professor Miguel Norbert-Ubarri (*University of Antwerp*) and his mother *doña* Anabel Ubarri for making sure I felt comfortable in San Juan.

In Santo Domingo, I could count on the support of Professor Lilia Ramos (*Pontífica Universidad Católica Madre y Maestra*, *Recinto Tomás de Acquino).* The assistance of Ginia Montes de Oca (*UNIBE Universidad Iberomericana*) was also of great value. Without their help, it would have been a lot more difficult to fill in the quota for young university-educated speakers.

While preparing the Havana fieldwork, I could enjoy the support of Sandra Valdés. Thanks to her help, I was able to get an academic visa. Through her friend Maia Barreda, Sandra also introduced me to Professor Alejandro Sánchez (*Universidad de La Habana, Facultad de Letras*). It was a relief to hear upon my arrival that Alejandro already had a list of potential participants in mind. During my stay, he continuously kept looking for more. I am also indebted with the staff of the *Casa de Abuelos Eterna Juventud* in the Centro Habana borough Cayo Hueso. Thanks to their collaboration I was able to interview older speakers at a very steady pace. Of course, this project would have failed without the help of my 72 participants, who in these busy times, were kind enough to dedicate an hour of their time to helping out a complete stranger.

Although my advisors Professor Frank Brisard (*University of Antwerp*) and Professor Nicole Delbecque (*University of Leuven*) only became involved in this project about a year ago, they have been of invaluable help. Their advice concerning theoretical and practical issues cannot be valued enough. Without their helpful criticism, the analyses would have been less rigorous and the conclusions drawn from them less solid. The

manuscript would also have contained many more typos and grammar mishaps, which they have tracked and corrected patiently.

By now, large parts of the data and analyses presented in Chapter 6 have been circulated in *Cognitive Linguistics*, *Language Variation and Change*, *Revista de Lingüística Iberoamericana*, and *Revista Española de Lingüística Aplicada* as well as at conferences in Canada, Europe, and the United States. Needless to say, this work has also benefited from the remarks of the anonymous reviewers these journals and venues assigned.

Last, but not certainly not least, I would like to thank my family, my girlfriend Sara, and my friends for their love and support. Without their continuous encouragement and loving understanding, I would most certainly have given up.

Jeroen Claes
Friday, October 31, 2014

Nederlandstalige samenvatting

In standaard Spaans geldt presentationeel *haber* (‘er is/er zijn’) als een onpersoonlijke constructie. Dit wil zeggen dat het werkwoord enkel de derde persoon enkelvoud aanneemt en dat het nominaal argument (e.g., *mangos* ‘mango’s’ in voorbeeld 1) zich als een lijdend voorwerp gedraagt. Er lijkt echter een groeiende tendens te zijn tot werkwoordscongruentie met de NP, zoals getoond in voorbeeld (2) (e.g., D'Aquino-Ruiz, 2008). Dit fenomeen staat bekend als de ‘pluralisering van *haber*’. Vanuit het perspectief van de Cognitieve Constructiegrammatica (e.g., Goldberg 1995, 2006a) zou dit kunnen betekenen dat het enkelvoudige constructieschema (<**AdvP** *haber* **Obj**>) afwisselend gebruikt wordt met een gepluraliseerd schema (<**AdvP** *haber* **Subj**>). De centrale hypothese die in dit proefschrift onderzocht wordt stelt dat in het Caribisch Spaans deze alternantie een taalverandering van onderen vormt. Tegen de achtergrond van deze stelling wordt de bijkomende hypothese onderzocht dat het fenomeen geconditioneerd wordt door drie algemene cognitieve factoren: gemarkeerdheid van de codering, statistische pre-emptie, en structurele *priming*.

(1) Sí, *había* mangos (SJ14H22 /SJ1672).
‘Ja, *er waren*$_{Sing}$ mango’s.’

(2) En mi época *habían* unos veinticinco, treinta alumnos por aula (LH01H22/LH17).
‘In mijn tijd *waren*$_{Plur}$ *er* een stuk of twintig, dertig leerlingen per klas.’

Om deze hypotheses te testen, voeren we een variationistische analyse uit. In drie recente steekproeven van het Spaans van Havana, Santo Domingo, en San Juan, gaan we de grammaticale en sociale distributie van de alternantie na, om vast te stellen of en in welke mate de variatie wijst op een taalverandering van onderen en een alternantie tussen de twee varianten van de presentationele *haber* constructie die geconditioneerd wordt door deze drie algemene cognitieve factoren.

Over het algemeen geven de resultaten aan dat sprekers van het Caribisch Spaans het werkwoord pluraliseren in gelijkaardige verhoudingen (Havana: 44.6%, N=934/2093; Santo Domingo: 46.7%, N=859/1841; San Juan: 41.4%, N=682/1649). De vergelijkende sociolinguïstische analyse (met gemengde-effecten logistische regressie, *conditional inference trees*, en *random forests*) toont weinig variatie als het gaat om grammaticale factoren. In de drie hoofdsteden hebben NP's die verwijzen naar typische actieketenhoofden, het gebruik van de meervoudige variant in het voorgaande discours

en alle werkwoordsvormen buiten de tegenwoordige en de verleden tijd een gunstig effect op de pluralisering. Het enige verschil bestaat erin dat in San Juan de afwezigheid van ontkenning een positief effect uitoefent op het gebruik van gepluraliseerd *haber*.

De associatie tussen meervoudig *haber* en sociale types varieert daarentegen sterker naargelang de taalgemeenschappen. In Havana wordt de alternantie gekoppeld aan lagere sociale klasse. In Santo Domingo, wordt *haber* pluralisering in verband gebracht met de middenklasse. In San Juan, ten slotte, wordt de pluralisering van *haber* geassocieerd met vrouwelijk gender. De resultaten suggereren ook een significante interactie tussen leeftijd en sociale klasse, in de zin dat jongere sprekers *haber* pluralisering lijken te koppelen aan lagere sociale klasse, terwijl het voor oudere sprekers juist een kenmerk van de middenklasse lijkt te zijn. Voor geen van de drie gemeenschappen kan een correlatie gevonden worden met de graad van formaliteit van het discours.

Deze resultaten suggereren dat de pluralisering van *haber* een geavanceerde taalverandering van onderen vormt (Labov, 2001: 308-309), die onderhevig is aan de drie algemene cognitieve factoren.

This page intentionally left blank

1

Introduction

The protagonist of this dissertation is the Spanish presentational construction with *haber* 'there is/there are'. In normative usage, this is an impersonal construction: it only takes the third-person singular verb ending and its nominal argument, *fiestas* 'parties' in example (1), behaves as a direct object. This is evident from the fact that it pronominalizes as an accusative pronoun in example (2) (Gili-Gaya, 1980: 78; Real Academia Española & Asociación de Academias de la Lengua Española, 2009: §41.6.b).

(1) Entonces, él siempre estaba velando en el periódico donde era que *había* fiestas (SJ03H22/SJ327).
'So, he was always watching in the newspaper where it was that *there were*$_{Sing}$ parties.'

(2) Interviewer: ¿Y también habían comidas que sólo se preparaban en fiestas, por ejemplo?
Participant: Sí, claro y todavía <u>las</u> *hay* (SD19M12/RD2547).
Interviewer: 'And were there also dishes that were only prepared on holidays, for example?'
Participant: [Yes, of course, and still <u>them</u>$_{Acc}$ *there are*$_{Sing}$.]
Participant: 'Yes, of course, and *there* still *are*$_{Sing}$.'

However, in many varieties of Spanish, including those spoken in Cuba, the Dominican Republic, and Puerto Rico (Vaquero, 1996: 44), presentational *haber* displays variable verb agreement with the NP (see example 3; e.g., D'Aquino-Ruiz, 2008; Kany, 1945/1951: 256-259). This phenomenon is known as the 'pluralization of (presentational) *haber'*.

(3) De seguro, no había televisión y, e, no *habían* computadores (SD04M22/RD437).
'Surely, there was no television and, er, *there weren't*$_{Plur}$ any computers.'

As will become clear in Chapter 2, it is already quite well known in which Spanish-speaking regions the alternation between pluralized and singular presentational *haber* occurs and which linguistic and social factors constrain it. Therefore, the goals of this investigation cannot be limited to obtaining yet another set of descriptions, this time for Cuban, Dominican, and Puerto Rican Spanish.

Drawing on Cognitive Construction Grammar (Goldberg, 1995, 2006), I will propose the main hypothesis that the variation between pluralized and singular presentational *haber* can be conceptualized as a linguistic change involving a competition between two variants of the presentational *haber* construction that are synonymous except for their associations with social groups.[1] Additionally, I will introduce the claim that this competition is not constrained by the type of highly specific linguistic factor groups that are usually proposed in variationist sociolinguistics, but rather by general cognitive constraints on linguistic expression. The comparative sociolinguistic analysis of the results obtained for Cuban, Dominican, and Puerto Rican Spanish will be a crucial element in making this point. Additionally, since language changes increase divergence between closely related varieties (Labov, 2010: 5; Wolfram & Schilling-Estes, 2003: 715), comparing the results for the three Caribbean dialects may also shed light on the question whether *haber* pluralization constitutes an ongoing linguistic change or, rather, an alternation that persists unchanged over time.

To achieve these goals, in Part A, I will start by providing the necessary background. Particularly, in Chapter 2, the literature on the pluralization of presentational *haber* will be reviewed. This will lead to the conclusion that *haber* pluralization occurs in many varieties of Spanish with similar linguistic constraints and recurring patterns of social covariation. However, previous dialectological and variationist work has not investigated the implications of these results for the nature of the constraints that condition *haber* pluralization.

Subsequently, Chapter 3 will introduce Cognitive Construction Grammar and the research questions. Drawing on the concepts provided by this cognitive-linguistic theory, I will formulate the main hypothesis and I will show that the trends documented in Chapter 2 can be explained as the reflexes of three general cognitive constraints on linguistic expression, namely, markedness of coding, statistical preemption, and structural priming. This will allow me to formulate a set of predictions regarding the frequency of *haber* pluralization in specific linguistic environments. Additionally, I will introduce Labov's (2001) Principles of Linguistic Change, which will lead me to hypothesize about the patterns of social covariation that *haber* pluralization may display.

Before continuing to test these predictions, Chapter 4 will introduce the comparative sociolinguistic methodology that was applied in this study. Particularly, I will introduce the sampling techniques, the fieldwork methods, the coding decisions, the statistical toolkit, and the way the data were compared.

[1] That is, their 'social meaning'.

Part B will present the results that were obtained in this investigation. Specifically, since the main hypothesis claims that *haber* pluralization involves a competition between two almost identical constructions, Chapter 5 will investigate the pragmatic, semantic, and syntactic properties of these two constructions. Particularly, it focuses on whether any differences in information status or conceptual-semantic meaning can be detected between pluralized and singular presentational *haber*. It will be shown that, except for the grammatical function of the NP argument, the two constructions are essentially identical. This suggests that the alternation can have social meaning.

Chapter 6 will test the hypotheses formulated in Chapter 3. Specifically, the focus in that chapter will be on evaluating whether the attested patterns of variation corroborate the claim that the alternation is conditioned by markedness of coding, statistical preemption, and structural priming. Chapter 6 will also try to establish with which social groups the frequent use of one of the variants is associated and whether this motivates seeing the variation as an ongoing linguistic change.

Drawing on these results, Chapter 7 will propose an account of how and when *haber* pluralization might have emerged. It first sketches a constructionist perspective on how singular presentational *haber* emerged in Late Latin. Then, I will show how singular presentational *haber* might have evolved into pluralized presentational *haber* in the context of the colonization of the (Spanish-Speaking) New World.

To conclude this dissertation, Chapter 8 will provide an overview of the results. Then, the research questions will be answered and these answers will be situated in a broader perspective in order to highlight the contributions of this investigation to both Cognitive Construction Grammar and variationist sociolinguistics.

This page intentionally left blank

Part A

Earlier studies, theory, and methods

This page intentionally left blank

2

The pluralization of presentational haber

In this chapter I will try to establish where the alternation between pluralized and singular presentational *haber* occurs and which linguistic and social factors constrain it. To this end, section 1 will review the (perceptual) dialectological literature on the pluralization of *haber*. Next, section 2 will provide an overview of the results of previous sociolinguistic analyses. The chapter concludes with a brief summary in section 3, which will highlight the main trends that emerge from the literature, as well as some of the limitations of earlier studies.

1. Earlier studies in (perceptual) dialectology

In general, *haber* pluralization has been documented in the majority of Spanish dialects (Real Academia Española & Asociación de Academias de la Lengua Española, 2009: §41.6b).[1] However, some differences can be found between regions when it comes to the degree to which speakers accept pluralized presentational *haber*. For Spain, dialectological research has shown that *haber* pluralization occurs occasionally in the varieties of Cantabria (Nuño-Álvarez, 1996: 190), Castilla la Vieja (Hernández-Alonso, 1996: 209), and Extremadura (Álvarez-Martínez, 1996: 180), but in these regions it is considered a substandard feature. Additionally, the pluralization of *haber* has been observed in Catalonia, eastern Andalucía, eastern Aragón, eastern Castilla La Mancha, eastern Murcia, and the Valencian Community (Gili-Gaya, 1980: 78; Maldonado de Guevara, 1980: 30; RAE & ASALE, 2009: §41.6b). In these areas, *haber* pluralization occurs in the speech of all social groups, and even in the written language, without any negative connotation attached to it (Blas-Arroyo, 1995-1996: 179, 1999: 55). By the same token, on the Canary Islands, pluralized *haber* can be found among all strata of the population, including university students and professors (Álvarez-Nazarío, 1991: 490; Catalán, 1989: 155, 199; Pérez-Martín, 2007). In contrast, Quilis's (1983: 94) study of agreement phenomena in spoken Madrid Spanish only reports two instances of pluralized *haber* against a total of more than 2,000

[1] Henceforth in this chapter, Real Academia Española & Asociación de Academias de la Lengua Española (2009) will be referred to as: RAE & ASALE (2009).

occurrences. These appear to have been mere slips of the tongue, because DeMello (1991: 449) does not find a single example for the varieties of Madrid and Seville.

Additionally, all Latin American varieties of Spanish feature *haber* pluralization, albeit in different proportions according to the local social evaluation (Fontanella de Weinberg, 1992a: 152-153; Moreno de Alba, 1995: 191). In this regard, Kany (1945/1951: 257-259) argues that the pluralization of *haber* occurs particularly frequently in Argentina, Central America, and Chile. In a review article of Kany (1945/1951), Flórez (1946: 379) adds that in Bogotá, presentational *haber* is also quite frequently pluralized by the lower and middle classes. In contrast, DeMello (1991: 449) shows that *haber* pluralization also occurs among university-educated speakers from this city. Moreover, the use of pluralized *haber* seems to be a feature of Latin American Educated Speech, since it appears in every city included in the *Proyecto del Estudio Coordinado de la Norma Lingüística Culta de las Principales Ciudades de Iberoamérica y de la Península Ibérica*, as can be seen in Table 2.1.

Table 2.1: *Haber* pluralization in the *Corpus of Latin American Educated Speech*: Numbers and percentages for pluralized *haber*

City	*N*	*%*
Buenos Aires	3/82	4%
Mexico City	7/92	8%
Bogotá	20/127	16%
Havana	12/45	27%
San Juan	29/95	31%
Caracas	55/153	36%
Santiago de Chile	51/132	39%
Lima	42/104	40%
La Paz	50/83	60%
Total	269/1038	26%

Source: DeMello (1991: 449)

Still, the distributions in the table suggest some differences across the continent. In particular, according to the data represented in Table 2.1, university-educated speakers from Caracas,[2] Havana, La Paz, Lima,[3] San Juan,[4] and Santiago de Chile pluralize *haber* in more than 25% of the cases. In contrast, Table 2.1 suggests that *haber* pluralization is rather infrequent among university-educated speakers from Bogotá, Buenos Aires, and Mexico City. However, at least for Buenos Aires and Mexico City, this is not corroborated by other studies, which have found *haber* pluralization to be a frequently occurring feature of educated Argentinean (Fontanella de Weinberg, 1987:

[2] See also Bentivoglio & Sedano (1996: 124).

[3] See also Caravedo (1996: 166).

[4] See also Vaquero (1996: 64).

154, 1992a: 152-153, 1992b: 36) and Mexican Spanish (Castillo-Trelles, 2007: 75; Lope-Blanch, 1996: 83; Montes de Oca, 1994: 21).

With respect to linguistic constraints, DeMello (1991: 460) finds that *haber* pluralization occurs frequently with the imperfect tense (*había, habían*). It appears to be less frequent with the preterit (*hubo*, *hubieron*) and, especially, the present tense (*hay, hayn*), for which DeMello (1991: 460) only documents the singular forms. Nevertheless, this does not imply that pluralized present- and preterit-tense tokens cannot be found. Quite on the contrary, a vernacular plural form has even emerged with the present tense, which is usually transcribed as *hayn*, *haen*, or *hain*. These forms, which are generally avoided by urban, educated speakers, have been documented in Antillean (Holmquist, 2008: 28; Vaquero, 1996: 64), Antioquian Colombian (Montes-Giraldo, 1982: 384), Argentinean (Kany, 1945/1951: 256-257), and Venezuelan Spanish (Lapesa, 1981: §133; Navarro-Correa, 1992: 98).

In order to investigate whether the low occurrence rate of *hubieron* is due to speakers associating this form with groups of low social prestige, Malaver (1999) compares the attitudes towards pluralized imperfect *habían* and pluralized preterit *hubieron* in Caracas, Venezuela. The results show that in Caracas, *habían* is generally considered to be correct. In contrast, pluralized preterit *hubieron* is considered as a feature of lower-class speech, although it is not uniformly judged to be incorrect (Malaver, 1999: 39-40). Freites-Barros (2003: 381, 2004: 43) replicates these results for San Cristóbal de los Andes. This suggests that in Venezuela, *hubieron* is a stigmatized form,[5] which is not accepted as part of the standard language. Freites-Barros (2003: 380, 2004: 41) also shows that formal education is a decisive factor in shaping speakers' attitudes towards *haber* pluralization, because university-educated speakers do not only disapprove more frequently of pluralized *haber,* they also motivate their choices by referring to rules of normative grammar.

Turning now to Caribbean Spanish, Vaquero (1996: 64) indicates that the Latin American tendency towards pluralizing *haber* can also be observed in the Antilles. Indeed, Kany (1945/1951: 259) cites examples for Cuba and Puerto Rico. For Cuba, Padrón (1949: 144) adds to this: "[i]n popular speech, the cases of verb agreement of the impersonal verb with its apparent subject are frequent."[6] However, judging from the data tabulated by DeMello (1991: 449; see Table 2.1), *haber* pluralization is not limited to the popular classes. This is also Domínguez-Hernández's (2007: 22) opinion.

[5] That is, a form associated with a group of low social prestige.

[6] In the original: "[e]n el habla popular son frecuentes los casos de concordancia del impersonal con el sujeto aparente" (Padrón, 1949: 144).

For the Dominican Republic, Henríquez-Ureña (1940/1982: 224) observes that in the Santo Domingo of the 1930s, only the lower classes use the pluralized construction. More recently, various authors (Alba, 2000: 23; Alvar-López, 2000: 338; González-Tapia, 1994: 94; Jiménez-Sabater, 1978: 178, 1984: 165) have argued that the pluralized forms are commonly heard throughout the country among all strata of the population. This is confirmed by Jorge-Morel (1978: 127), who finds that, in Santo Domingo, individuals of all educational backgrounds report using the pluralized preterit form *hubieron*, although more uneducated participants admit to using it.[7] Additionally, Fernández's (1982: 93, 102) attitude study shows that two thirds of the students of the *Pontífica Universidad Católica Madre y Maestra*[8] consider future- and imperfect-tense pluralized *haber* to be correct and part of the standard language.[9] However, the pluralized preterit form *hubieron* is considered correct by less than 50% of the students (46%, N=62/135 students). Yet, more recently, Alvar-López (2000: 338) finds that "[a]ll social classes use *hubieron*".[10] This is confirmed by Alba (2004: 323), who reports that 53% (N=73/138 students) of his sample of university students consider this form to be correct. These data appear to warrant the conclusion that, contrary to the stigmatization of *hubieron* observed among Venezuelans, the majority of the Dominicans is firmly convinced of the normative correctness of pluralized *haber*, as is argued by Alba (2004: 28).

For Puerto Rico, Navarro-Tomás (1948: 131) observes that pluralized instances of presentational *haber* "are not only heard in rural settings, but also, as is the case in other countries, in the informal language of the urban classes."[11] Similarly, Álvarez-Nazario (1991: 490, 709) points out that *haber* pluralization can be found in the speech of all social classes. Indeed, Table 2.1 suggests that university-educated speakers do not refrain from using pluralized *haber*. Vaquero's (1978: 135-140) attitude study points in the same direction, as it shows that about one third of the students of the Río Piedras campus of the *Universidad de Puerto Rico* identify pluralized imperfect *habían* as correct (34%, N=98/288 students). As was the case in the Dominican Republic, a similar figure is found for pluralized preterit *hubieron* (29%, N=84/288

[7] Jorge-Morel (1978: 127) reports that 15% (N=3/20) of her university-educated participants respond using pluralized *haber*. In contrast, 56.5% (N=13/23) of the participants with junior or senior secondary education state that they use pluralized *haber*. 81.5% (N=22/27) of the group including analphabets and participants with only primary education respond using pluralized *haber*.

[8] A university located in Santiago de los Caballeros, the second largest city of the Dominican Republic.

[9] Fernández (1982: 102) indicates that the pluralized morphological future form *habrán* is rated as correct by 64.4% (N=87/135) of the students. Pluralized imperfect *habían* is approved by 61.5% (N=83/135) of the students.

[10] In the original: "[t]odas las clases sociales emplean *hubieron*" (Alvar-López 2000: 338).

[11] In the original: "se oyen … no sólo en los medios rurales sino también, como en otros países, en el lenguaje familiar de las clases urbanas" (Navarro-Tomás, 1948: 131).

students). Finally, López-Morales (1992: 147) reports that 63% of university-educated speakers residing in San Juan consider pluralized imperfect *habían* to be correct.

In sum, this section has shown that *haber* pluralization constitutes a wide-spread phenomenon that appears in many varieties of Spanish. In Caribbean Spanish, it occurs in the speech of all social strata, including university-educated speakers. In Venezuela, the use of pluralized preterit *hubieron* is stigmatized, but in Cuba, the Dominican Republic, and Puerto Rico speakers appear to consider all non-present pluralized forms of presentational *haber* part of the standard language. In contrast, pluralized present-tense *hayn* is absent from the speech of university-educated speakers, but it may still be found in Antillean Spanish.

2. Earlier studies in variationist sociolinguistics

The pluralization of presentational *haber* has been investigated in Canarian, Mexican, Puerto Rican, Salvadorian, and Venezuelan Spanish. In this section, the results of these studies will be reviewed in chronological order by country, which will allow me to identify recurring linguistic constraints and patterns of social covariation. The latter are of particular interest if one wishes to establish whether the phenomenon constitutes an ongoing language change 'from above', 'from below' or rather a 'stable variable'. Before turning to the review of the literature, in the following section, these types of sociolinguistic variability and their characteristic patterns of social covariation will be introduced.

2.1 Symptoms of change and stability

Language changes from below are spontaneous linguistic evolutions that emerge in the middle class (Labov, 1966/2006: 206, 2001: 188) and spread upward through the social hierarchy below speakers' level of consciousness (Labov, 1966/2006: 206-207, 1972: 179). As this type of language change occurs without speakers realizing it, changes from below have a high probability of going to completion (Labov, 1972: 178-180, 2001: 517-518). In situations of change from below, female speakers (Labov, 2001: 292), middle-class speakers (Labov, 2001: 188), and younger speakers (Labov, 1994:43-72) use the innovative forms more frequently. Additionally, the rates of use of the innovative variant do not decrease when formality rises (Labov, 1972: 239, 2001: Chap. 3; Silva-Corvalán, 2001: 248-249), that is, when the amount of attention that is paid to speech increases (Labov, 1966/2006: 59-86, 1972: 99).

Eventually, speakers may grow aware of a change from below, in which case it may become stigmatized. In that case, a conscious effort (through education, mass media, and other linguistic institutions) may be made to replace the innovative variant with a form, usually borrowed from another variety or language, that is judged more favorably (Labov, 1994: 78). Such often retrograde changes, typical of standardization

processes, are known as 'changes from above' (Labov, 1972: 179). In situations of change from above, female speakers and younger speakers typically use the stigmatized variant less often and its frequency is a monotonic function of formality and speakers' social class (Labov, 1966/2006: 204-206, 2001: Chap. 3). Thirdly, stable variables represent a pattern of social covariation similar to that of changes from above, with one difference: they do not covary with speakers' age (Labov, 2001: 119).

2.2 Venezuela

Turning now to earlier studies of *haber* pluralization carried out in Venezuela, Bentivoglio & Sedano (1989: 64) base their findings on the *Corpus of Educated Speech* of Caracas and a sample of 70 sociolinguistic interviews with speakers of Caracas Spanish, stratified by age, gender, and social class. With these data, Bentivoglio & Sedano (1989: 67) show that speakers use the pluralized forms in 52.1% (N=88/169) of the cases. Additionally, these authors investigate the effects of age (15-29 years, 30-45 years), gender (female, male), social class (lower class, middle class, upper class), and the three linguistic factor groups exemplified in Table 2.2. In their univariate analyses, age, gender, reference of the NP, social class, and verb tense turn out to be statistically significant. For the last factor group, Bentivoglio & Sedano (1989: 72) show that *haber* pluralization occurs most often with the imperfect tense. As for the reference of the NP, they observe that human-reference NPs favor pluralized *haber* (Bentivoglio & Sedano, 1989: 73). Regarding social factors, pluralized *haber* occurs more often in the speech of lower-class, male, and middle-class speakers (Bentivoglio & Sedano, 1989: 67). Additionally, young speakers of these groups pluralize presentational *haber* more often, which may suggest a change in progress.

Table 2.2: Linguistic factor groups considered by Bentivoglio & Sedano (1989)

Factor groups	*Examples*
Reference of the NP	
Human	*Había* muchos norteamericanos. '*There were*$_{Sing}$ a lot of North Americans.'
Nonhuman	*Hubo* reuniones en la comunidad. '*There were*$_{Sing}$ meetings in the community.'
Reinforcement of the idea of plurality	
Not reinforced	*Hubo* reuniones en la comunidad. '*There were*$_{Sing}$ meetings in the community.'
Reinforced by indefinite quantifiers	*Había* muchos norteamericanos. '*There were*$_{Sing}$ many North Americans.'
Reinforced by coordination of nouns	En el fondo *habían* conchas de mango, pepas de mango, gorros de baño. 'At the bottom, *there were*$_{Plur}$ mango shells, mango seeds, bathing caps.'
Reinforced by placing the NP before *haber*	Hay exceso de gente para las comodidades que *había*. 'There is an excess of people for the commodities that *there were*$_{Sing}$.'
Verb tense	

Source: Examples taken from Bentivoglio & Sedano (1989: 65-66)

Subsequent studies on Venezuelan Spanish have refined and replicated Bentivoglio & Sedano's (1989) results. Specifically, Domínguez, Guzmán, Moros, Pabón & Vilaín (1998: 26) analyze a sample of 48 sociolinguistic interviews with speakers of the Spanish of Mérida, stratified by age, gender, and social class. These authors find that, in this variety, presentational *haber* is pluralized in 61.4% (N=58/94) of the cases (Domínguez et al., 1998: 29). Domínguez et al. (1998: 29-32) proceed to analyze the same variables as Bentivoglio & Sedano (1989), with identical results. However, their frequency counts should be handled with great care, because they are based on a very limited amount of occurrences.[12] These authors also submit a multiple-choice sentence completion questionnaire to 25 students of the *Universidad de Los Andes*, who were instructed to complete the test utterances with one of the suggested forms and to motivate their responses. The data obtained with this instrument suggest that young speakers of the variety of Mérida prefer to use pluralized *haber,* which they generally consider to be the correct variant (Domínguez et al., 1998: 32-35).

Díaz-Campos (1999-2000), on the other hand, compares a stratified sample of Caracas Spanish with the *Corpus of Educated Speech of Santiago de Chile* also analyzed by

[12] Less than two cases per speaker.

DeMello (1991). Generally speaking, the results of this investigation show that both cities display similar rates of *haber* pluralization (Santiago de Chile: 26%, N=12/46; Caracas: 25%, N=9/36). This study also corroborates, at least for Caracas, Bentivoglio & Sedano's (1989) results concerning the effects of human-reference NPs and the imperfect tense (Díaz-Campos, 1999-2000: 224-225). For Santiago, however, none of the linguistic factors is statistically significant (Díaz-Campos, 1999-2000: 224-225). Additionally, Díaz-Campos (1999-2000: 226) shows that in both communities, women favor the pluralized forms.

Later, Díaz-Campos (2003: 4-5) selects a sample of 96 sociolinguistic interviews from the *Estudio Sociolingüístico de Caracas* corpus, stratified by age, gender, and social class. Since these data were collected ten years after Bentivoglio & Sedano's (1989) datasets,[13] this restudy of Caracas may shed some initial light on the real-time development of *haber* pluralization. In general, Díaz-Campos (2003: 8) shows that speakers pluralize presentational *haber* in 54.3% of the cases (N=245/451). Additionally, Díaz-Campos (2003: 5-7) performs a fixed-effects logistic regression analysis with six factor groups: age (14-29 years, 30-45 years, 46-60 years, 61+ years), gender (female, male), social class (lower class, middle class, upper class), and the three linguistic factor groups specified in Table 2.3. The results show that the variation is only constrained by the speaker's social class and the verb tense. For the first of these factors, like Bentivoglio & Sedano (1989), Díaz-Campos (2003: 8) finds that lower- and middle-class speakers favor pluralized *haber*. Concerning the verb tense, Díaz-Campos (2003: 8) observes pluralized presentational *haber* more often with both the imperfect (*había, habían*) and the present perfect tense (*ha habido, han habido*). In contrast, the preterit (*hubo, hubieron*) and the other tenses disfavor pluralization. Regarding the question as to whether the alternation constitutes an ongoing linguistic change from below or rather a stable variation, Díaz-Campos (2003: 9) notes that, although the frequencies of pluralized presentational *haber* have increased in those ten years, the phenomenon has hardly spread from the imperfect paradigm to others, with the exception of the present perfect. Moreover, the fact that age does not seem to be a factor in this variation is also suggestive of a stable variable, at least for these two tenses (Díaz-Campos, 2003: 11). However, as Díaz-Campos (2003: 11) observes, resolving this matter will require investigating the effect of formality on *haber* pluralization, as the frequency of stable variables typically declines when formality rises (Labov, 2001: Chap. 3). Therefore, he calls for investigations that "address this issue by observing the interaction between the pluralization of *haber* and factors such as speech style, social class, sex, and age" (Díaz-Campos, 2003: 11).

[13] That is, in 1987.

Table 2.3: Linguistic factor groups considered by Díaz-Campos (2003)

Factor groups	*Examples*
Reference of the NP	
Human	*Habían* profesores.
	'*There were*$_{Plur}$ professors.'
Nonhuman	No *había* edificios.
	'*There weren't*$_{Sing}$ buildings.'
Reinforcement of the idea of plurality	
Not reinforced	*Habían* profesores.
	'*There were*$_{Plur}$ professors.'
Reinforced by an adjective	*Habían* buenos proyectos.
	'*There were*$_{Plur}$ good projects.'
Reinforced by coordination of nouns	*Habían* hornos de cal, alfarería y cuestiones.
	'*There were*$_{Plur}$ lime ovens, pottery, and things.'
Reinforced by a determiner	*Habían* otros grupos.
	'*There were*$_{Plur}$ other groups.'
Reinforced by a determiner and an adjective	*Habían* unos árboles grandes.
	'*There were*$_{Plur}$ some big trees.'
Verb tense	

Source: Examples taken from Díaz-Campos (2003: 5-6)

D'Aquino-Ruiz (2004) studies a larger sample from the *Estudio Sociolingüístico de Caracas* corpus also analyzed by Díaz-Campos (2003).[14] With these data, D'Aquino-Ruiz (2004: 18) shows that speakers pluralize *haber* in 63% of the cases (N=477/754). Additionally, she investigates nine linguistic (see Table 2.4 and Table 2.5) and three social factor groups: age (14-29 years, 30-45 years, 46-60 years, 61+ years), gender (female, male), and social class (lower class, lower middle class, middle class, upper middle class, upper class). Of these, only four turn out to be significant in the fixed-effects logistic regression analysis: absence/presence of negation, social class, type of plural noun, and verb tense. For the first factor group, D'Aquino-Ruiz (2004: 18) finds that pluralized presentational *haber* occurs less often in clauses involving negation. For social class, the analysis indicates that lower-class speakers pluralize *haber* more often (D'Aquino-Ruiz, 2004: 18). Regarding the type of plural noun, D'Aquino-Ruiz (2004: 18) shows that singular mass nouns, either specific or unspecific, disfavor pluralized *haber*. For the verb tense, synthetic tenses[15] are shown to favor pluralized presentational *haber*, whereas the compound tenses and the preterit disfavor

[14] D'Aquino-Ruiz (2004: 5) analyzes the interviews of 160 participants, whereas Díaz-Campos (2003) only considers the interviews of 96 speakers.

[15] 'Synthetic' refers to the forms of *haber* that consist of just one word, as opposed to the compound tenses and to expressions in which *haber* forms a verb phrase with an aspectual or modal auxiliary.

pluralization (D'Aquino-Ruiz, 2004: 18). Finally, D'Aquino-Ruiz (2004: 16) argues that, as age and gender do not seem to condition the variation, *haber* pluralization is most adequately described in terms of a stable variation. However, the alternations between pluralized and singular *haber* do appear to be spreading from the lower to the upper classes, which leads her to conclude that the phenomenon could become a future change from below (D'Aquino-Ruiz, 2004: 22).

Table 2.4: Linguistic factor groups considered by D'Aquino-Ruiz (2004)

Factor groups	*Examples*
Absence/presence of aspectual or modal auxiliaries	
Absent	*Había* muchas peleas entre salones.
	'*There were*$_{Sing}$ many fights between classrooms.'
Present	Mínimo *deben haber* dos personas de acuerdo.
	'Minimally, *there have to be*$_{Plur}$ two people who agree.'
Absence/presence of negation	
Absent	*Tienen que haber* productos superfluos también.
	'*There* also *have to be*$_{Plur}$ superfluous products.'
Present	Nunca *hubo* zapateros.
	'*There were*$_{Sing}$ never shoemakers.'
Definiteness of the NP	
Definite	Allí *habían* el partido comunista y el MIR.
	'Over there, *there were*$_{Plur}$ the communist party and the MIR.'
Indefinite	También *habían* fiestas de quince años con Billos.
	'*There were*$_{Plur}$ also fifteenth-birthday celebrations with Billos.'
Reference of the NP	
Human	Donde *habían* equis cantidad de estudiantes.
	'Where *there were*$_{Plur}$ x amount of students.'
Nonhuman	Siempre *habían* muchos choques.
	'*There were*$_{Plur}$ always many clashes.'
Type of clause	
All others	No *habían* abastos sino pulperías.
	'*There weren't*$_{Plur}$ supermarkets, but grocery stores.'
Relative clause	Todas las matas de mango que *habían* aquí.
	'All the mango trees that *there were*$_{Plur}$ here.'

Source: Examples taken from D'Aquino-Ruiz (2004: 5-13)

Table 2.5: Linguistic factor groups considered by D'Aquino-Ruiz (2004) (continuation)

Factor group	*Examples*
Type of NP	
Implicit NP	…que *había* dos. '…that *there were*$_{Sing}$ two.'
Lexical and pronominal NP	Se observan algunas canchas que antes no las *habían*. [Are observed some courts that before them$_{Acc}$ *there weren't*$_{Plur}$.] 'Some courts are observed that *there weren't*$_{Plur}$ before.'
Lexical NP	No *hubo* problemas. '*There weren't*$_{Sing}$ problems.'
Pronoun	Los *había* preciosos. [Them$_{Acc}$ *there were*$_{Sing}$ beautiful.] '*There were*$_{Sing}$ beautiful ones.'
Type of plural noun	
Mass noun	Creo que en Letras *había* un grupito también. 'I think that in Arts *there was*$_{Sing}$ a little group as well.'
Plural count noun	*Había* noches que yo no dormía. '*There were*$_{Sing}$ nights that I didn't sleep.'
Specific mass noun	Y *había* un grupito ya grande de muchachos. 'And *there was*$_{Sing}$ quite a large group of kids already.'
Word order	
Haber + NP	O *había* pequeñas manifestaciones. 'Or *there were*$_{Sing}$ small manifestations.'
Implicit NP	En el año *habían* muchas. 'Throughout the year, *there were*$_{Plur}$ many.'
NP + *haber*	Diferentes de esas fragatas que *habían* aquí. 'Different from those frigates that *there were*$_{Plur}$ here.'
Verb tense	
Compound tenses	No me acuerdo, así, que *haya habido.* 'I don't remember, like that, that *there have been*$_{Sing}$.'
Synthetic tenses	De repente, *habrán* otras cosas. 'Suddenly, *there will be*$_{Plur}$ other things.'
Preterit	Y parece que *hubieron* muertos. 'And it appears that *there were*$_{Plur}$ casualties.'

Source: Examples taken from D'Aquino-Ruiz (2004: 5-13)

To gain more insight into the question as to whether *haber* pluralization is most accurately described as a change in progress or as a stable variable, D'Aquino-Ruiz (2008: 118) compares the results obtained by DeMello (1991) and Bentivoglio & Sedano (1989) with data taken from her 2004 study. This database shows that lower-

to middle-class individuals and younger speakers consistently use pluralized presentational *haber* most often (D'Aquino-Ruiz, 2008: 116-118). Additionally, the overall rates of *haber* pluralization increase from 39% (total N=74) in Bentivoglio & Sedano's (1989) upper-class data, recorded in 1972-1974, to 66% (N=465/706) in the *Estudio Sociolingüístico de Caracas* data, gathered in 1987 (D'Aquino-Ruiz, 2008: 118-119). These findings, together with the fact that the phenomenon does not appear to be stigmatized, lead D'Aquino-Ruiz (2008: 120-121) to conclude that the alternation between pluralized and singular presentational *haber* constitutes a linguistic change from below nearing completion.

Finally, Freites-Barros (2008) analyzes a sample of 128 interviews with residents of San Cristóbal de Los Andes, stratified by age, gender, and regional origin. In general, pluralized *haber* appears in no less than 82% (N=245/298) of the cases.[16] Freites-Barros (2008: 44-47) further examines five linguistic (see Table 2.6 and Table 2.7) and three social factor groups: gender (female, male), age (15-30 years, 31-45 years, 46-60 years, 60+ years), and regional origin (rural, urban). The fixed-effects logistic regression analysis withholds three of these, namely, reference of the NP, reinforcement of the idea of plurality, and type of NP (Freites-Barros, 2008: 53). For the first two, Freites-Barros (2008: 53) finds that pluralized *haber* is favored with human-reference NPs and the presence of elements that reinforce the idea of plurality. For the third factor, the results suggest that implicit NPs and pronouns disfavor pluralized presentational *haber* (Freites-Barros, 2008: 53).

Table 2.6: Linguistic factor groups considered by Freites-Barros (2008)

Factor group	*Examples*
Reference of the NP	
Human	Porque en ese entonces sí *habían* profesores que valía la pena lo que enseñaban. 'Because in that time, *there were*$_{Plur}$ professors that it was worth what they taught.'
Nonhuman	A veces, cuando *habían* vacas, se ordeñaban las vacas. 'Sometimes, when *there were*$_{Sing}$ cows, the cows were milked.'

Source: Examples taken from Freites-Barros (2008: 44-47)

[16] These include third-person presentational *haber* in all non-present tenses and first-person plural *haber* in all tenses (Freites-Barros, 2008: 47).

Table 2.7: Linguistic factor groups considered by Freites-Barros (2008) (continuation)

Factor group	*Examples*
Reinforcement of the idea of plurality	
Not reinforced	No *habían* vagones para transportar cargamento pesado. '*There weren't*$_{Plur}$ wagons to transport heavy loads.'
Reinforced by indefinite quantifiers, numerals, or coordination of nouns	*Hubieron* muchos animales muertos. '*There were*$_{Plur}$ many dead animals.'
Specificity	
Nonspecific	Siempre que *habían* velorios yo me escapaba. 'Whenever *there were*$_{Plur}$ wakes, I escaped.'
Specific	En esa casa *habían* dos viejitas que vestían bien. 'In that house, *there were*$_{Plur}$ two little old ladies who dressed well.'
Type of NP	
Implicit NP	Yo iba a buscarle sus cigarrillos. Si en la bodega que estaba más cerca no *habían* tenía que ir a la otra bodega. 'I fetched his cigarettes. If in the shop that was closest *there weren't*$_{Plur}$, I had to go to the other shop.'
Lexical NP	*Habían* trapiches pa' moler caña. '*There were*$_{Plur}$ sugar cane mills to grind sugar cane.'
Pronoun	La gente aquí toda es buena; los *habrá* por allá, pa' otros barrios, pero aquí no. [All the people around here are good; them$_{Acc}$ *there will be*$_{Sing}$ around there, towards other boroughs, but not here.] 'All the people around here are good; *there will be*$_{Sing}$ around there, towards other boroughs, but not here.'
Verb tense	
Word order	
Haber + NP	Sí, *había* torturas. 'Yes, *there were*$_{Sing}$ tortures.'
Implicit NP	No sé qué hacían con él, pero lo cierto es que no *habían.* 'I don't know what they did with him, but the sure thing is that *there weren't*$_{Plur}$.'
NP + *haber*	Ladrones, no *habían*. 'Thieves, *there weren't*$_{Plur}$.'

Source: Examples taken from Freites-Barros (2008: 44-47)

2.3 Mexico

Using a written, open-ended sentence completion questionnaire, Montes de Oca (1994: 17-18) investigates the use of pluralized and singular presentational *haber* among 120 middle- and upper-class participants from Mexico City, stratified by age, gender, and social class. She tests the influence of five linguistic factor groups: absence/presence of an aspectual or modal auxiliary (e.g., *poder* 'can', *deber* 'must', *ir a* 'be going to', *tener que* 'have to'), absence/presence of quantifiers (absent, present), absence/presence of a relative clause in the NP (absent, present), reference of the NP (human, nonhuman), and verb tense (aspectual or modal auxiliary + *haber*, imperfect, present perfect, preterit, subjunctive present). Montes de Oca (1994) also considers three social factor groups: age (15-25 years, 26-35 years, 36 to 55 years), gender (female, male), and social class (middle class, upper class). In general, the results show that Mexicans pluralize presentational *haber* in 30.3% of the cases (Montes de Oca, 1994: 21). For the linguistic factor groups, Montes de Oca's (1994: 26-33) univariate analyses suggest that human-reference NPs and the presence of quantifiers and relative clauses favor pluralized *haber*. Additionally, for the verb tense, she finds that the presence of aspectual or modal auxiliaries favors pluralization, whereas synthetic and compound tenses disfavor it (Montes de Oca, 1994: 22-25). Turning to social factors, Montes de Oca (1994: 21) observes that her upper-class participants pluralize presentational *haber* more often (35.5%, N=469/1320) than her middle-class participants (24.5%, N=324/1320). In order to test whether speakers also pluralize *haber* with pronominal nominal arguments, Montes de Oca (1994: 34-35) administers a multiple-choice sentence completion questionnaire to a sample of 30 upper-class participants between 15 and 35 years of age. The results of this follow-up study suggest that, although verb agreement occurs infrequently with accusative pronouns (10.8%, N=13/120 responses), it is not at all impossible (Montes de Oca, 1994: 34-35).

Castillo-Trelles (2007: 75-76), in turn, analyzes a sample of recording sessions with 24 speakers of the variety of Mérida, Yucatán, stratified by age, data collection method, educational achievement, gender, and knowledge of Maya. In general, the verb is pluralized in 53% (N=85/160) of the cases. Castillo-Trelles (2007: 80-81) also tests three linguistic (see Table 2.8) and four social factor groups: age (7-18 and 55-73 years, 19-54 years), data collection method (interview, word combination task),[17] educational achievement (primary, middle school, high school or more), and knowledge of Maya (bilingual Maya-Spanish, monolingual in Spanish). The fixed-effects logistic regression analysis retains only the absence/presence of quantifiers and gender as significant factor groups for the recording sessions. The results show that the absence of quantifiers and female gender favor pluralized *haber*. Castillo-Trelles (2007: 76-77) also administers a multiple-choice sentence completion questionnaire to

[17]After the interview, the participants are instructed to combine snippets of sentences using the verb *haber*.

another group of 54 speakers. In their responses, the verb is pluralized in 55% of the cases (N=518/936). For this dataset, the results of Castillo-Trelles's (2007: 81) fixed-effects regression analysis point to significant effects of age, gender, knowledge of Maya, reference of the NP, and word order. For age, the questionnaire data show that the age group comprising youngsters and retirees pluralizes presentational *haber* most often. Regarding gender and knowledge of Maya, the results suggest that women and bilinguals favor pluralization. Finally, for the reference of the NP and word order, the regression analysis shows that human-reference NPs and the word order NP + *haber* have positive effects on *haber* pluralization.

Table 2.8: Linguistic factor groups considered by Castillo-Trelles (2007)

Factor group	*Examples*
Absence/presence of quantifiers	
Absent	Antes *hubieron* perros callejeros. 'Before, *there were*$_{Plur}$ stray dogs.'
Present	Antes *habían* menos perros callejeros. 'Before *there were*$_{Plur}$ less stray dogs.'
Reference of the NP	
Human	Antes *habían* menos personas. 'Before, *there were*$_{Plur}$ less people.'
Nonhuman	Antes *había* menos robos. 'Before, *there were*$_{Plur}$ less burglaries.'
Verb tense	
Aspectual or modal auxiliary + *haber* and compound tenses	En el futuro, si seguimos así, *podría haber* más perros callejeros. 'In the future, if we continue like this, *there could be*$_{Sing}$ more stray dogs.'
Imperfect	Antes no *habían* semáforos. 'Before, *there weren't*$_{Plur}$ traffic lights.'
Morphological future	En el futuro *habrán* más fraccionamientos. 'In the future, *there will be*$_{Plur}$ more urbanizations.'
Word order	
Haber + NP	Dentro de unos años *habrán* fraccionamientos subterráneos. 'Within a few years, *there will be*$_{Plur}$ underground urbanizations.'
NP + *haber*	Antes, los fraccionamientos no *habrán*. 'Before, the urbanizations *there won't be*$_{Plur}$.'

Source: Examples taken from Castillo-Trelles (2007: 77-78)

2.4 El Salvador

Quintanilla-Aguilar (2009: Chap. 4.2.1) analyzes a sample of 48 interviews with native speakers of San Salvador Spanish, stratified by age, educational achievement, and gender. In general, he finds that pluralized presentational *haber* occurs in 79.6% (N=218/274) of the cases (Quintanilla-Aguilar, 2009: 146).[18] In addition, Quintanilla-Aguilar (2009: 126-129) investigates the influence of ten factor groups: the six linguistic factor groups listed in Table 2.9 and four social factor groups, namely, age (18-35 years, 50+ years), discourse spontaneity (elicited by a question containing singular presentational *haber,* spontaneous), educational achievement (basic secondary or less, university), and gender (female, male). Of these, Quintanilla-Aguilar's (2009:172-173) fixed-effects logistic regression analysis only retains two: discourse spontaneity and verb tense. For the first factor group, the regression analysis suggests that speakers are less likely to use pluralized *haber* in answers to questions containing singular presentational *haber*. In Quintanilla-Aguilar's (2009: 162) analysis, this shows that *haber* pluralization is less frequent in formal registers.[19] For the second factor group, Quintanilla-Aguilar (2009: 173) finds that the imperfect tense favors pluralized presentational *haber*, whereas all other tenses disfavor pluralization. Finally, the fact that age and gender did not turn out to be significant constraints on the variation lead him to conclude that, in San Salvador, *haber* pluralization is most adequately described in terms of a stable variable (Quintanilla-Aguilar, 2009: 180).

[18] These include third-person presentational *haber* in all non-present tenses and first-person plural *haber* in all tenses (Quintanilla-Aguilar, 2009: 153).

[19] These results may also suggest that *haber* pluralization is subject to structural priming, as I will argue in Chapter 3.3.2.3.

Table 2.9: Linguistic factor groups considered by Quintanilla-Aguilar (2009)

Factor group	*Examples*
Absence/presence of quantifiers	
Absent	Porque allá no *habían* escuelas en donde estudiar. 'Because over there, *there weren't*$_{Plur}$ schools to study in.'
Present	Durante la guerra *hubieron* más de setenta mil personas muertas. 'During the war, *there were*$_{Plur}$ more than seventy thousand people dead.'
Absence/presence of negation	
Absent, present	No examples provided by Quintanilla-Aguilar (2009: 129-130).
Reference of the NP	
Human	Siempre *van a haber* pobres. '*There are* always *going to be*$_{Plur}$ poor people.'
Nonhuman	No *habían* pugnas entre los sindicatos y directivas. '*There weren't*$_{Plur}$ conflicts between the unions and directives.'
Type of verb phrase	
Aspectual or modal auxiliary + *haber*	El Señor siempre dijo: "Pobres siempre *van a haber.*" 'The Lord always said: "Poor people *there will* always *be*$_{Plur}$."'
Compound form	Porque ahí cuando *han habido* terremotos y todo eso no se han caído casas ni nada. 'Because over there, when *there have been*$_{Plur}$ earthquakes and all this stuff, houses have not fallen down or anything.'
Synthetic form	Allí *habían* ya prostíbulos con mucha evidencia. 'Over there, *there were*$_{Plur}$ already unconcealed brothels.'
Verb tense	
Imperfect, other tenses, preterit	No examples provided by Quintanilla-Aguilar (2009: 128-129).
Word order	
Haber + NP	Yo tengo entendido de que hubieron muchísimos más muertos. 'I understand that *there were*$_{Plur}$ many more casualties.'
NP + *haber*	¡Bastantes muertos *hubieron*! 'Enough casualties *there were*$_{Plur}$!'

Source: Examples taken from Quintanilla-Aguilar (2009: 128-130)

2.5 Canary Islands

Pérez-Martín (2007: 505) studies the pluralization of presentational *haber* in the Canarian variety of Spanish spoken on the island El Hierro. To this end, she analyzes a sample of 56 ten-minute interviews, stratified by age, gender, and social class. As this corpus provides a very small number of tokens for other tenses than the imperfect, Pérez-Martín (2007: 508) restricts the analysis to the alternation between *había* and *habían*. The results show that speakers pluralize imperfect *haber* in 76.7% of the cases (N=46/60). Additionally, she analyzes the effect of two linguistic factor groups:[20] reference of the NP (human, nonhuman) and reinforcement of the idea of plurality (not reinforced, reinforced by placing the NP in preverbal position or by including a quantifier in the NP) (Pérez-Martín, 2007: 509). Pérez-Martín (2007: 509-510) also evaluates the influence of three social factors: age (20-34 years, 35-54 years, 54+ years), gender, (female, male), and social class (lower class, middle class, upper class). Regarding the linguistic factor groups, Pérez-Martín's (2007: 509) univariate analyses suggest that Canarian speakers pluralize *haber* more often with human-reference NPs and NP arguments that reinforce the notion of plurality. For the social factors, the univariate analyses suggest that *haber* pluralization correlates with middle-aged, male, and upper-class speakers. However, it should be stressed that these results are to be interpreted with great care, because they are based on a very limited number of tokens[21] and, therefore, might reflect data imbalance rather than patterns of social covariation.

2.6 Puerto Rico

In a series of three recent articles, Esther Brown and Javier Rivas (Brown & Rivas, 2012; Rivas & Brown, 2012, 2013) analyze a corpus of Caguas, Cayey, and San Juan Spanish and the *Corpus of Educated Speech of San Juan* also used by DeMello (1991). The results show that speakers pluralize the verb in, respectively, 44% (N=83/190; Brown & Rivas, 2012: 329) and 58% (N=41/98; Rivas & Brown, 2013: 110) of the cases and that the choice between pluralized and singular *haber* is controlled by the properties of the NP argument and the verb tense. Specifically, Brown & Rivas (2012: 331) argue that nouns that are predominantly used as subjects in Spanish trigger the reanalysis of the NP slot of the presentational construction with *haber* more often. In their view, this suggests that the mental representations of nouns include a 'grammatical relation probability'[22] and that this probability leads speakers to interpret the NP slot of the presentational *haber* construction either as a subject or as an object.

[20] Pérez-Martín (2007) does not provide examples.

[21] Less than two cases per speaker.

[22] In other words, Brown & Rivas (2012) claim that speakers store the frequency with which a particular noun is used in a specific grammatical function and that this probability determines the likelihood that they will use pluralized presentational *haber*. In Chapter 6.2.1, we will get back on this proposal.

In a related paper (Rivas & Brown, 2012), these authors explore the hypothesis that *haber* pluralization is constrained by the semantic contrast between 'stage-level' and 'individual-level' nouns. As Rivas & Brown (2012: 74) observe, the categories individual level and stage level were originally formulated as a way to capture the semantic differences between predicates that denote permanent, intrinsic properties of entities ('individual-level predicates'; e.g., *smart*) and those that describe more transient characteristics ('stage-level predicates'; e.g., *cold*). As this distinction was devised to categorize predicates, the nouns that occur with presentational *haber* cannot readily be classified as being individual-level or stage-level. Therefore, the sense given to these notions by Rivas & Brown (2012) deviates significantly from their original use. Particularly, these authors code nouns such as *elecciones* 'elections' in example (1), which refer to events or entities "that have an understood beginning and ending" (Brown & Rivas, 2012: 81) as 'stage-level nouns'. In turn, nouns such as *directores* 'directors' and *superintendentes* 'superintendents' in example (2), which "have a preferential interpretation as beginning prior to and continuing past the point of reference of the predication" (Rivas & Brown, 2012: 81), were coded as 'individual-level nouns'.

(1) Porque fue cuando *hubo*, este, las elecciones (from Rivas & Brown, 2012: 81).
'Because it was when *there were*$_{Sing}$, er, the elections.'

(2) Pero también *habían* directores, superintendentes (from Rivas & Brown, 2012: 82).
'But *there were*$_{Plur}$ directors, superintendents.'

The results show that individual-level nouns favor pluralized *haber*. Additionally, Rivas & Brown (2012: 84, 2013: 111) also observe that pluralized presentational *haber* occurs most frequently with the imperfect tense.[23] Let us summarize now the main trends that emerge from the literature.

[23] In Chapter 6.2.1.1 we will get back on this analysis.

3. Summary: Trends and limitations

In this chapter, I have reviewed the dialectological and the sociolinguistic literature on the pluralization of presentational *haber*. This has shown that *haber* pluralization constitutes a widely diffused alternation that appears to exist in Canarian, Latin American, and Peninsular Spanish. Most of the sociolinguistic studies that were reviewed in this chapter indicate that pluralized *haber* is more frequent when the NP has human reference and the verb is conjugated in the imperfect tense, a compound tense or forms a verb phrase with an aspectual or modal auxiliary. The results of D'Aquino-Ruiz (2004) and Quintanilla-Aguilar (2009) also suggest that *haber* pluralization is sensitive to the absence/presence of negation. Yet, none of the studies that were reviewed above provides an analysis of the phenomenon that goes beyond describing the effect of specific, isolated linguistic factors.

When it comes to patterns of social covariation, the frequent use of pluralized *haber* appears to correlate with lower social class in Venezuela. This is especially true for the pluralized preterit form *hubieron* (Freites-Barros, 2003, 2004; Malaver, 1999). However, this does not seem to be the case in the Dominican Republic and Puerto Rico, where *hubieron* is considered correct by large segments of the population, including university students (Alba, 2004; Fernández, 1982; Vaquero, 1978). In addition, the pluralization of *haber* does not seem to correlate with age, which could indicate that we are dealing with a very slowly progressing linguistic change from below (D'Aquino-Ruiz, 2008; Díaz-Campos, 2003; Fontanella de Weinberg, 1992b: 44) or with a stable variable (Quintanilla-Aguilar, 2009). However, the results for gender are not as consistent. That is, Bentivoglio & Sedano's (1989) results, which show that male speakers favor *haber* pluralization, could be interpreted as suggesting a situation of stable variation (Labov, 2001: Chap. 3). In contrast, the fact that Díaz-Campos (1999-2000) and Castillo-Trelles (2007) find that women pluralize *haber* more often points towards a language change from below (Labov, 2001: 292). As differences between gender groups tend to become smaller or even disappear in later stages of language changes from below (Labov, 2001: 308-309), Freites-Barros's (2008) and Quintanilla-Aguilar's (2009) results, which document no signs of gender differentiation, may also point in this direction. The fact that Castillo-Trelles's (2007) results suggest that speakers do not pluralize *haber* less often when performing a word combination task, which explicitly focuses all attention on language, is a further argument in favor of considering the variation as a change from below. Yet, as Díaz-Campos (2003: 11) observes, only a systematic evaluation of the interaction between age, gender, social class, and formality will allow us to shed new light on this issue.

There are other reasons to study an alternation that has already been well documented. First, I have not been able to find investigations that take structural priming[24] (e.g., Pickering & Ferreira, 2008) into account.[25] The tendency to recycle structures that have appeared in previous discourse has proven to be a potent constraint on multiple syntactic alternations (Labov, 1994: 566-570; Martín-Butragueño, 1999: 231-232; Pereira-Scherre & Naro, 1992; Travis, 2005, 2007; Weiner & Labov, 1983: 49-56). Second, apart from Díaz-Campos (1999-2000) and DeMello (1991) – both based on limited and outdated datasets – I have not been able to find investigations that compare two (or more) Latin American dialects. Yet, comparing two or more varieties could shed new light on the issue whether *haber* pluralization is most accurately conceptualized as a stable variable or rather as a slowly progressing linguistic change, because the latter inevitably leads to divergence among closely related varieties (Labov, 2010: 5; Wolfram & Schilling-Estes, 2003: 713). At the same time, such a comparison could also further our insight into the cognitive factors that constrain *haber* pluralization (and linguistic variation in general), its social distribution, and the origin of the alternation. Finally, exploring how all of this can be modeled in Cognitive Construction Grammar (Goldberg, 1995, 2006a) may offer refreshing insights into the way this cognitive-linguistic theory can contribute to the study of syntactic variation and change. This will be the topic of the next chapter.

[24] That is, the influence of immediately preceding plural verb forms.

[25] Quintanilla-Aguilar (2009) does investigate some limited priming effects, under the guise of his variable 'discourse spontaneity'. However, he does not take into account the effect of the speaker's earlier mentions of a particular variant of *haber*.

3

Cognitive Construction Grammar, research questions, and hypotheses

In this chapter, I will be concerned with the theoretical background, the research questions, and the hypotheses of this study. Section 1 introduces Cognitive Construction Grammar, including the way sociolinguistic variation and language change can be modeled in this theory. In section 2, the research questions will be presented. Subsequently, section 3 addresses the hypotheses that will be pursued in Chapter 6. The chapter concludes with a brief summary in section 4.

1. Cognitive Construction Grammar

In this section, I will present a thumbnail sketch of Cognitive Construction Grammar. It should be clear from the onset that the aim is not to provide a comprehensive introduction, as this can already be found in, among others, Croft (2007), Croft & Cruse (2004: Chap. 9-10), and Goldberg (2003, 2009, 2010). Particularly, section 1.1 defines the basic concepts Cognitive Construction Grammar hinges upon. Section 1.2 is concerned with the usage-based character of construction grammar(s). Section 1.3, in turn, presents arguments in favor of a construction-based approach to language. Section 1.4 focuses on the meaning of constructions and section 1.5 identifies some constraints on their use. In section 1.6, the typical formalism of Cognitive Construction Grammar is introduced. Section 1.7 demonstrates that Cognitive Construction Grammar can easily accommodate sociolinguistic variation. Finally, section 1.8 discusses the way language change and, particularly, form-function reanalysis can be conceptualized in this framework.

1.1 Basic concepts

The first concept that should be introduced is that of ‘grammatical construction’. In (Cognitive) Construction Grammar, this notion indicates a conventional pairing of form and meaning (Croft, 2007: 472; Goldberg, 1995: 4, 2009: 94; Lakoff, 1987: 467). This includes words, but also abstract patterns that involve more than one entity and may contain variables. By way of illustration, Table 3.1 represents an overview of constructions of different sizes and degrees of schematicity.

Table 3.1: Overview of constructions of different sizes and degrees of schematicity

Construction type	*Examples*
Word	*tentacle*, *gangster*, *the*
Partially filled word	<*post*-N>, <V-*ing*>
Complex word	*textbook*, *drive-in*
Idiom (filled)	*like a bat out of hell*
Idiom (partially filled)	<*believe* one's *ears*/*eyes*>
Covariational conditional	<The Xer the Yer>
	The more you watch the less you know
Ditransitive	<**Subj** V **Obj1 Obj2**>
	She gave him a kiss.
Passive	<**Subj** auxiliary VP (PP_{by})>
	The cell phone tower was struck by lightning.

Source: Adapted from Goldberg (2009: 94)

The table shows that grammatical constructions can be taken to represent every aspect of language structure (Croft & Cruse, 2004: 254) and that no principled distinction is assumed between the nature of generalizations, such as, for example, argument-structure, and lexical items. Both are conceptualized as form-function pairings, which only differ in schematicity (Langacker, 1987: 28, 1990: 16). In other words, Cognitive Construction Grammar, like other cognitive-linguistic theories, pictures the grammar as a network of form/meaning pairs ranging from morphemes to abstract construction schemata (Croft, 2007: 471; Fillmore, 1989: 34; Langacker, 1987: 36, 1990: 12, 2007: 427).[1]

Since language is claimed to consist entirely out of units that specify both form and meaning, constructionist approaches are nonmodular (Croft, 2007: 471). As a result, there can be neither strictly syntactic elements[2] nor underlying forms of representation (Goldberg, 2006a: Chap. 2, 2009: 97; Langacker, 1987: 22-27, 1990: 18-19, 2010: 109). As the network of constructions is considered to provide speakers with conventional symbolic resources for the encoding of conceptualizations, constructionist approaches are also nongenerative, nonconstructive, and nonderivational (Langacker, 1990: 15). Rather, constructionist theories claim that speakers apply these symbolic resources in language using nothing but domain-independent, general cognitive abilities such as categorization (Langacker, 1990: 15-16).

[1] That is, as a lexicon/grammar continuum.

[2] That is, meaningless forms.

As a 'usage-based' model, Cognitive Construction Grammar also recognizes that a substantial part of speakers' day-to-day language use consists of ready-made units. Therefore, the actual computation of novel expressions is argued to be fairly limited (Bybee, 2001: 15, 2006: 713; Langacker, 1987: 58-59). In the next section, this approach to language will be considered in more detail.

1.2 Usage-based linguistics

Cognitive Construction Grammar, like all cognitive-linguistic theories, maintains that language structure is determined by usage (Bybee, 2001, 2006: 730; Bybee & Beckner, 2010; Langacker, 1987: 57, 1990: Chap. 10, 2008: 457-458). This approach to language is known as 'usage-based'. One of its most important implications is that language is assumed to represent variability and change at all points in time (Langacker, 1991: 369). Additionally, usage-based models emphasize the importance of a quantitative approach to the study of language structure, because each time a construction or a lexical item is used or processed, it is believed to become gradually more accessible (Bybee, 2001: 6-7, 2003b, 2006: 715; Bybee & Beckner 2010: 829; Goldberg, 2006a: 12, 2009: 98; Langacker, 1990: Chap. 10, 2007: 425). As we will see in section 1.7 and section 1.8, because of these characteristics, cognitive-linguistic theories are particularly well-suited to model language variation and change (Clark, 2007, 2008: 269-270).

The claim that frequency affects mental representation is evident in the assumed effect of 'token frequency', that is, the number of times a structure occurs in discourse. High token frequency has a conserving effect, in the sense that it converts a compositional expression into a single item and reinforces its representation; this is called 'entrenchment' (Bybee, 2001: Chap. 5, 2003a: 153, 2003b: 604-619, 2006, 2007: 960-962; Bybee & Beckner, 2010: 840; Goldberg, 1995: 79; Langacker, 1987: 59-60, 1991: 48). In turn, because entrenched expressions can be accessed easily, they are "preferentially produced over items that are licensed but are represented more abstractly, as long as the items share the same semantic and pragmatic constraints" (Goldberg, 2006a: 94).[3] This general cognitive constraint is known as 'statistical preemption'.

[3] See also Langacker (1990: 285-286).

1.3 Arguments for constructions

One of the advantages of assuming a usage-based, construction-based grammar architecture is that we can account fairly easily for the process of grammaticalization (Bybee, 2009: 347), that is, for the process that converts fully compositional expressions into grammatical morphemes (e.g., Croft, 2000: Chap. 6). In addition, construction grammars handle 'peripheral' syntactic phenomena, such as idioms[4] and information-focus driven alternations [5] with the same ease as 'core' syntactic phenomena, such as, for example, transitivity. However, it is probably fair to say that construction grammar and, especially, Cognitive Construction Grammar, has been shown to be most useful in the study of argument-structure alternations.

In this regard, recall that in generative syntax, the verb is considered to be the main determinant of clause-structure (Chomsky, 1995: 238). However, this is seriously challenged by two of its implications. First, portraying the verb as the pivot of the clause implies that for every alternation that is uncovered, a copy of the verb must be stored in the lexicon (Goldberg, 2001: 504). This would mean that even for the very limited sample of *to kick* expressions in example (1), we would need seven copies of the verb to accommodate the fluctuations in clause structure and meaning. Second, if this role were to befall to the verb, we would be forced to claim that verbal neologisms, like *to flubber* in example (2), are stored in the lexicon with their argument-structure specifications (Goldberg, 2009: 95). In contrast, in Cognitive Construction Grammar, the overarching argument-structure construction is assumed to determine the overall meaning of the clause, the number of arguments, and their argument role. This way, with only one verb entry we can account for the full range of variation (Goldberg, 1995: 9-13, 16, 1999: 198).

(1) a. I *kick* the bed with my heel (Davies, 2008-, Fiction).
b. They run past crumbling homes and *kick* balls *into* goals with no nets (Davies, 2008-, Press).
c. My mother and her friends talk in low voices while the men roll up the hose and *kick* the shards of glass *off* the driveway (Davies, 2008-, Fiction).
d. It doesn't matter that those black people are big and fierce, when it comes to fighting we can *kick the shit out of* them (Davies, 2008-, Magazine).
e. But I don't *kick at* them (Davies, 2008-, Fiction).
f. The rookie Mario Bates kick-started the New Orleans running attack, then Morten Andersen *kicked* them *to* a home victory (Davies, 2008-, Press).
g. His fingers fastened on something damp and cool and resilient. *It kicked* (Davies, 2008-, Fiction).

[4] See, for example, Fillmore, Kay & O'Connor (1988), Goldberg (1996), Israel (1996), Jackendoff (1997, 2008), and Kay & Fillmore (1999).
[5] See, for example, Michaelis & Lambrecht (1996), Goldberg (2001), and Goldberg & Ackerman (2001).

(2) That thing just *flubbered* into my room (constructed example).

Of course, the fact that a theoretical construct provides parsimonious solutions to long-standing problems in linguistic theory (e.g., Goldberg, 1995: Chap. 1) does not necessarily mean that it is psychologically adequate. However, a series of studies conducted by Goldberg and others have provided strong psycholinguistic arguments in favor of argument-structure constructions. Specifically, Goldberg (2006b: 417) has shown that listeners are able to guess the meaning of a nonsense verb correctly based upon the construction pattern it occurs in. This was already evident from the example with *to flubber*. By the same token, experiments reported by Goldberg (2006a: 112, 2009: 95) suggest that argument-structure constructions are better indicators for the meanings of expressions than individual verbs. Still other experiments indicate that listeners rely on the meaning of argument-structure constructions to determine the verb's sense when they are confronted with novel noun-to-verb extensions (Goldberg, 2009: 95).[6] However, the most compelling evidence of the psychological reality of argument-structure constructions probably consists in the fact that "individual abstract constructions can be distinguished using fMRI data, even when content, open-class words, complexity, and frequency are held constant" (Allen, Pereira, Botvinick & Goldberg, 2012: 178).

1.4 The meaning of constructions

Like all cognitive-linguistic theories, Cognitive Construction Grammar assumes that the meaning of a lexical unit is taken to comprise "everything speakers know about the type of entity designated" (Langacker, 2007: 432),[7] including a set of background assumptions. This approach to semantics is known as 'frame semantics', in which the notion of 'frame' is taken to denote

> a system of categories structured in accordance with some motivating context. … The motivating context is some body of understandings, some pattern of practices, or some history of social institutions, against which we find intelligible the creation of a particular category in the history of the language community (Fillmore, 2006: 381).

This definition also indicates that frames consist of two parts: what I have called 'background assumptions' are usually referred to as 'background frame' (Goldberg, 2010: 40) or 'base' (Langacker, 1987: 180-189). The foreground, in turn, is most commonly indicated with the term 'profile' (Langacker, 1987: 189).

Turning now to the meaning of verbs, Cognitive Construction Grammar proposes that this class of lexical items refers to conceptualizations of specific events. Since an event

[6] For example, when listeners are confronted with the novel extension of *crutch* to *to crutch* in a ditransitive construction, they interpret *He crutched her the ball* as implying that someone used a crutch to transfer the ball to someone else. In contrast, when listeners are presented with *He crutched her*, they take this transitive expression to denote that someone used a crutch to hit someone else (Goldberg, 2009: 95).

[7] See also Croft & Cruse (2004: 30) and Langacker (1987: 154-166).

presupposes entities participating in it, the frames of verbs specify how many participants partake in the event and what role they fulfill in it. This is expressed in the form of verb-specific 'participant roles' (see example 3).[8]

(3) Participant roles: hit <**hitter**, **hittee**>.

Besides listing the participants in the event and their roles, the frames of verbs also specify which participants are profiled (Goldberg, 2005b: 225).[9]

Argument-structure constructions, in turn, are assumed to refer to conceptualizations of event types rather than to specific events. Specifically, Langacker (1991: 294-295) proposes that grammatical constructions encode conceptual archetypes. Goldberg (1995: Chap. 2.3.5), for her part, argues that "constructions designate scenes essential to human experience" (Goldberg, 1995: 39). Still, as both of these proposals involve abstraction over and idealization of observed events of the same type, they are compatible with Lakoff's (1987: 489-490) claim that constructions encode 'Idealized Cognitive Models' (ICMs, henceforth) of events. According to Cienki (2007),

> ICMs are proposed as a way in which we organize knowledge, not as a direct reflection of an objective state of affairs in the world, but according to certain cognitive structuring principles. The models are idealized, in that they involve an abstraction, through perceptual and conceptual processes, from the complexities of the physical world. At the same time, these processes impart organizing structure—for example, in the form of conceptual categories (Cienki, 2007: 176).

Since argument-structure constructions designate abstractions over events of the same type, they also assign more abstract roles to the participants partaking in them. These are labeled 'argument roles' (e.g., *agent*, *patient*, *receiver*…).[10] Similarly, like verbs, argument-structure constructions also specify which participants are profiled. As will become evident in the next section, this is a major constraint on the co-occurrence of verbs and constructions.

---

[8] These roles are event-specific instances of more general argument roles such as, for example, agent, patient, receiver, etc.

[9] This appears to be a major source of semantic variation between verbs. Consider, for example, *to rob* and *to steal*. Both verbs denote that something is taken from someone without permission, but they profile different portions of the event frame. That is, *to rob* profiles the thief and the victim, whereas *to steal* profiles the thief and the stolen goods (Goldberg, 1995: 45). As a matter of convention, profiled participants are marked with boldface font (Goldberg, 1995: 45).

[10] However, Cognitive Construction Grammar does not propose a finite list of argument roles. Rather, these follow directly from the construction's basic sense and, hence, "are more specific and numerous than traditional thematic roles" (Goldberg, 2005a: 23). In other words, labels such as *agent, patient,* etc., should be interpreted as mere shorthand notations that capture the semantic characteristics associated with the slots of constructions (Goldberg, 2005b: 224).

1.5 Constraints on verbs and constructions

Although cognitive linguistics credits speakers with a virtually limitless creative potential, this does not mean that just any verb can be used with just any argument-structure construction. Rather, it appears that the meanings of verbs and constructions should at least be relatable to each other in some fashion, as can be deduced from the oddness of example (4).

(4) *Little Johnny slept the pineapples from the ceiling (constructed example).

In the most prototypical cases, the event denoted by the verb instantiates the ICM designated by the argument-structure construction (Goldberg, 2010: 53). This relationship is illustrated in example (5), where *to hand* refers to a specific case of the CAUSE-RECEIVE ICM designated by the ditransitive construction. In these cases, the contribution of the verb to the meaning of the overall expression is limited, as it only adds more specific information (Goldberg, 1995: 51).

(5) I *handed* her the reins, while she glanced at me below the brim of her hat (Davies, 2008-, Fiction).

Frequently, however, the verb refers to an event that is not of the type designated by the argument-structure construction. Still, there are multiple ways in which events of two different types can relate to one another. For our purposes, it suffices to introduce here only the precondition relationship, shown in example (6).[11] In this example, the event of baking a cake constitutes a logical precondition for the sense of transfer denoted by the ditransitive construction.

(6) It's my grandmother's 90th birthday and I wanted to *bake her a cake* (Davies, 2008-, Fiction).

The participants of verbs and argument-structure constructions impose additional constraints. Specifically, the construction and the verb should share at least one participant (Goldberg, 1995: 65), which should satisfy the following two principles:

> *The Semantic Coherence Principle:* The participant role of the verb and the argument role of the construction must be semantically compatible. In particular, the more specific participant role of the verb must be construable as an instance of the more general argument role. General categorization processes are responsible for this categorization task and it is always operative.
>
> *The Correspondence Principle*: The semantically salient profiled participant roles are encoded by grammatical relations that provide them a sufficient degree of discourse prominence, i.e., by profiled argument roles. An exception arises if a verb has three argument roles; in this case, one can be represented by an unprofiled argument role (and realized as an oblique argument).

[11] See Goldberg (1995: 60-66, 2010: 53) for other, possible relationships between verbs and constructions.

> The Correspondence Principle can be overridden by specifications of particular constructions (Goldberg, 2005b: 225-226, emphasis added).[12]

Whenever these conditions are met with, verbs and argument-structure constructions can co-occur freely (Goldberg, 2006a: 10, 2009: 96).[13] Let us now consider the typical formalism of Cognitive Construction Grammar.

1.6 The formalism

In any discussion of the formalism of Cognitive Construction Grammar it should be clear from the beginning that the theory treats its typical notation merely as a device that helps exposition and discussion, without making any claims of psychological reality in its regard (Goldberg, 2006a: Chap. 10.4). With this reservation, the framework uses box diagrams to depict speakers' full grasp of their language.[14] Therefore, semantic, syntactic, and, when relevant, pragmatic constraints on the use of constructions are depicted in the diagrams.

As an example, let us consider how such a box diagram would look like for the English ditransitive construction CAUSE-RECEIVE, exemplified in (7).

(7) *I hand him his water* and he pushes north (Davies, 2008-, Press).

The participant roles of the verb *to hand* are listed in (8).

(8) Participant roles: hand <**hander, handed, handee**> (Goldberg, 1995: 51).

The composite structure of the ditransitive and *hand* is represented in Figure 3.1. In the diagram, *Sem* indicates the semantic pole of the construction, with the small capitals representing the CAUSE-RECEIVE ICM. The profiled argument roles, which need to be fused with the verb's participant roles, are listed on the right hand side of the ICM. The arrows specify which argument roles are instantiated by which participant roles and how these are mapped onto syntactic functions. Next to the line connecting CAUSE-RECEIVE with *hand*, a letter *R* appears. This letter indicates the type of relation that holds between the event denoted by the verb and the ICM designated by the argument-structure construction (Goldberg, 1995: 50-51, 1996: 40-41, 2005b: 228). The following section will explore how sociolinguistic variation can be modeled in Cognitive Construction Grammar.

---

[12]The Correspondence Principle is considered to be valid crosslinguistically "insofar as lexically profiled roles are expressed by core grammatical relations when they are expressed" (Goldberg, 2005b: 235-236).

[13] The correspondence principle is a default principle, which ensures that in normal, declarative expressions, lexical semantics, syntax, and discourse pragmatics are aligned (Goldberg, 2005a: 36-37). In more complex expressions, violations of this principle can be expected (Goldberg, 2005a: 25, 2005b: 226-227).

[14] According to Langacker (1990), "[t]he ultimate goal of linguistic description is to characterize, in a cognitively realistic fashion, those structures and abilities that constitute a speaker's grasp of linguistic convention" (Langacker, 1990: 15).

Figure 3.1: The English ditransitive construction instantiated by *to hand*

Sem	CAUSE-RECEIVE	<	**agent**	**recipient**	**patient**	>
R: instance,	|R		|	|	|	
means	*hand*	<	**hander**	**handee**	**handed**	>
	↓		↓	↓	↓	
Syn	V		Subj	Obj1	Obj2	

Source: Adapted from Goldberg (1995: 51)

1.7 Cognitive Construction Grammar and sociolinguistic variation

The study of language variation forms an integral part of Cognitive Construction Grammar. This is corroborated by the attention paid by Goldberg and her collaborators to argument-structure variation.[15] Cognitive Construction Grammar (and cognitive linguistics generally) is also interested in vernacular usage, because

> [i]f speakers use grammatical patterns that a speech community (through its normative grammars) does not readily embrace, then the combined facts that such patterns (a) are used, and (b) have not been (explicitly) taught, guarantee the importance of such structures in language; it is not an indication of their triviality. When we encounter forms that we have not been explicitly taught – not to mention expressions that speakers are warned (by prescriptive grammars) against using – we know that we are touching on something very basic, something that must be rooted in our cognitive behavior independently of what others have attempted to impose on us (Fried & Östman, 2004: 15-16).

Recent work in variationist sociolinguistics suggests that speakers actively use such vernacular features to signal social subgroup membership or to take stances[16] (Clark, 2007: 9-10, 2008: 269-270; Eckert, 2008; Kiesling, 2005, 2009). However, this is not to be conceived as all-or-nothing choices between discrete forms. Rather, speakers add layers of social meaning to the propositional content of their messages by actively and agentively controlling (Chambers, 2009: 266; Clark, 2008: 269; Hudson, 1996: 246) the frequencies of the alternating forms they use (Labov, 1966/2006: 84-85, 208, 2010: 372; Weinreich, Labov, & Herzog, 1968: 163). Of course, alternations will only pick up social meaning if the variants encode the same propositional content, that is, refer to the exact same conceptualization of a thing, event, or event type (Lavandera, 1978, 1984: Chap. 1).

In cognitive linguistics, social meaning can be thought of as a byproduct of the domain-general process of categorization (e.g., Lakoff, 1987: Chap. 1). That is, we are continuously and unconsciously seeking out the similarities and the differences between the people and the objects around us. This leads us to classify our

[15] See section 1.3.

[16] That is, attitudes towards the interlocutor and/or the propositional content of the message.

surroundings into types.[17] In this process, we first observe that certain social types of people exist. Subsequently, we also observe that individuals instantiating a certain social type use particular variable forms more often than others (Clark, 2007: 9-10, 2008: 269-270).[18] This, in turn, leads us to establish a metonymic link between (knowledge on) the social type and the distributions of linguistic variants (Kristiansen, 2008: 67-68). If this social type co-occurs frequently with these distributions in a variety of situations, this link may become entrenched in the speech community. In that case, its members conventionally associate the social type with the variable use of particular linguistic forms.[19] As a result, the alternation acquires social meaning (Clark, 2007: 9-10; Eckert, 2009: 14-15).

In this light, it is rather unproblematic to suppose that if two argument-structure constructions refer to exactly the same ICM, each will include a social distribution specification. This specification ensures that the use of one of these argument-structure constructions at a certain rate will be interpreted as signaling social subgroup membership ('1st-order index' in Silverstein's, 2003 terms) and, potentially, everything the participants of a usage event can be expected to know about this group ('1+nth order index' in Silverstein's, 2003 terms),[20] such as, for example, stances often taken by its members. Of course, all the knowledge speakers possess on a particular social type will not become activated every time the relevant distribution is observed. Rather, the context of the usage event will activate or background certain things language users know about a social group (Eckert, 2009: 14-15), just like this context constrains the potential meanings of constructions (Langacker, 1987:154-166). This explains how the multiple context-dependent interpretations of sociolinguistic variables arise (e.g., Bucholtz, 2009; Eckert, 2008; Kiesling, 2005, 2009).[21]

---

[17] Nevertheless, this is not to say that categorization is a conscious process. Rather, it is a corollary of the way our memories work. That is, when we are presented with an array of instances of the same phenomenon, we automatically abstract what is common to them and retain that commonality, whereas the particulars of each instance slowly decay (Bybee, 2001: 28; Goldberg, 2006a: 62; Langacker, 2007: 425).

[18] See also Bybee (2001: 29), Bybee & Beckner (2010: 830, 846), Geeraerts (2006: 27-28), Goldberg (2006a: 10), and Langacker (1987: 63, 2008: 42, 2010: 90-91).

[19] Experimental data suggest that his is the case. First, attitude research has shown that judges are able to identify social types and to deliver statements on them from exposure to the distribution of just one sociolinguistic variable (e.g., Freites-Barros, 2004; Mack, 2011; Malaver, 1999; Preston, 2002). Second, Hay, Warren, & Drager (2006) have shown experimentally that listeners draw on the perceived social identity of the speaker (e.g., working-class vs. middle-class attire) to disambiguate isolated near-homophones created by the New Zealand *near-square* vowel merger (e.g., *air – ear*, *bear – beer* are no longer minimal pairs; Hay et al., 2006: 462). This would be impossible if the distributions of individual linguistic features were not metonymically linked with speakers' mental representations of social types.

[20] See, for example, Langacker (2007: 432).

[21] Eckert (2008), for example, writes: "the meanings of variables are not precise or fixed but rather constitute a field of potential meanings – an indexical field, or constellation of ideologically related meanings, any one of which can be activated in the situated use of the variable" (Eckert, 2008: 454). In other words, the same variable

In sum, the discussion in this section has shown that Cognitive Construction Grammar can readily accommodate the social meaning of alternating constructions that refer to the same event type. However, in this study I will only be concerned with the links Cubans, Dominicans, and Puerto Ricans establish between the distribution of pluralized and singular presentational *haber* and (knowledge on) social types. Let us consider now how language change can be modeled in this framework.

1.8 Cognitive Construction Grammar and language change

As was already observed in section 1.2, usage-based theories claim that language structure is determined by usage. In that section, we have also seen that these theories recognize that language represents variability and change at every moment in time. In this section, one type of language change will be discussed, namely, form-function reanalysis.

At face value, linguistic change appears to be paradoxical, because "language, as an instrument of communication, would work best if it did not change at all" (Labov, 2001: 5).[22] Still, it has long been recognized that linguistic innovations occur on a day-to-day basis, although these rarely go on to constitute linguistic changes (Croft, 2000: 118; Kerswill, 1996: 178). Indeed, cognitive linguists consider that speakers innovate continuously, because "the targets of categorization are consistently more specific and elaborate than the categorizing structures and very often deviate from their specifications" (Langacker, 2008: 459). As a result, "[l]anguage use is always pushing the envelope of established convention" (Langacker, 2008: 459). Still, the basic idea is that speakers try to adhere to the norms of their speech community (Croft, 2000: 118). However, paradoxically, this may actually induce them to break convention (Croft, 2000: 118). One such case is 'form-function reanalysis', the unconscious remapping of the form-function relationships of (parts of) constructions.

Yet, before we can turn to way speakers establish new form-function mappings in language use, first, it might be useful to consider the way these arise in language acquisition. In this regard, cognitive approaches assume that children first get a hold on what particular expressions mean in specific situations (Croft & Cruse 2004: 323; Tomasello, 1995: 151, 2005: 186, 2006: 443-447, 2007: 1099-1103). When their vocabulary grows, they are confronted with the need of categorizing the learned input into classes (Goldberg, Casenhiser, & Sethuraman, 2004: 303-304). This, in turn, leads children to abduct[23] an argument-structure construction that is common to the

---

may signal, for example, subgroup membership in one context, a stance in another, or a completely novel social meaning in yet another context.

[22] See also Labov (1982: 23).

[23] According to Andersen (1973), abduction is a type of reasoning that takes a result, then applies a known or tentative generalization to the result to obtain an inference about something. For example, "given the fact that

exemplars. Subsequently, this construction schema can be used in novel ways (Goldberg, 1999: 206-209; Goldberg et al., 2004: 305; Tomasello, 1995: 151, 2005: 193, 2006: 449, 2007: 1103-1107).

Form-function reanalysis, then, consists in reapplying this unconscious abductive reasoning to established form-function mappings. Particularly, while speaking, language users unconsciously contrast the form of constructions with that of others that have similar meanings and are used in similar situations. This leads them to detect a majority pattern (say, places are encoded as obliques). As a result, speakers infer that the syntactic form of, for example, the place they wish to encode is that of an oblique, even though the construction schema may normally encode it as another grammatical function (Croft, 2000: 118-119, 141).

While this shows how speakers may break linguistic convention without wanting to do so (Croft, 2000: 118), we are still left with two questions. First, what counts as the start of a linguistic change? And, second, why do some of these online reanalyses become adopted throughout the speech community, whereas most of them do not? For Weinreich et al. (1968: 186), the answer to the first question is clear: the 'actuation'[24] of a linguistic change coincides with the adoption of the innovation by (some social groups in) the speech community.[25] The answer to the second question, however, is usually referred to as the 'actuation problem' (Weinreich et al., 1968: 186; Labov, 1972: 283), precisely because it is somewhat more complicated.

In this regard, innovative forms tend to occur sporadically and erratically in speech communities during a very long period of time (Croft, 2000: 60). Therefore, the exact moment at which a change takes off and its sociohistorical characteristics are impossible to pinpoint in hindsight (Israel, 1996: 226; Labov 1972: 277). Still, research in sociolinguistics suggests that the actuation of linguistic changes is usually triggered by the arrival of important numbers of adult language or dialect learners to the community and the ensuing koineization (Labov, 1982: 82-83, 2001: 314, 2010: Chap. 5).[26] As a result of the process described in the previous section, once a reformed pattern has become entrenched in a particular group, it becomes associated with this particular social type and those who wish to identify themselves with it (Chambers, 2009: 266; Croft, 2000: 181). From there on, the fate of the change depends on the sociological profiles of the individuals that introduce the change to the community at large.

Socrates is dead, we may relate this fact to the general law that all men are mortal and guess that Socrates was a man" (Andersen, 1973: 775).

[24] Or, in other words, the start of a linguistic change.

[25] See also Labov (1972: 277, 1982: 46).

[26] See Chapter 7.2 for discussion of koineization.

Specifically, if the change is transmitted through young, middle-class women of the second or third generation of immigrant ethnic groups who are in overt disagreement with dominant norms, have close ties to their own neighborhoods, but still frequent individuals of different social statuses in and beyond the own local neighborhood (Labov, 2001: Part C), it will constitute a change 'from below'. If working class males introduce the innovative form, it will most likely become a 'stable variable'. [27] Let us now consider the research questions.

2. Research questions

Against the backdrop of the discussion in the previous sections and taking into account the points raised in the final section of Chapter 2, this study will focus on the following questions:

I. Cognitive factors in *haber* pluralization
   - What are the cognitive factors that constrain the pluralization of presentational *haber* in Caribbean Spanish?
   - How can these constraints be modeled in Cognitive Construction Grammar?

II. Social factors in *haber* pluralization
   - What is the social distribution of the pluralization of presentational *haber* in Havana, Santo Domingo, and San Juan?
   - Do these distributions justify the characterization of the phenomenon as a linguistic change from below?

III. Comparison of the Caribbean dialects
   - What are the differences (if any) between the social distributions and the effect of the cognitive constraints as they are observed in the varieties of Havana, Santo Domingo, and San Juan?
   - What do these results indicate about the emergence of *haber* pluralization and the nature of the constraints that condition it?

In the following section, tentative answers to these questions will be formulated, beginning with the list of cognitive factors that may potentially constrain *haber* pluralization.

[27] See Chapter 2.2.1 for discussion of change from above, change from below, and stable variation.

3. Hypotheses

As observed Chapter 2.3, earlier variationist studies have typically found that *haber* pluralization is constrained by the reference of the nominal argument and the verb tense. Additionally, the literature suggests that the differentiated use of pluralized and singular presentational *haber* signals social class membership and, in certain speech communities, gender identity. These patterns of social covariation point in the direction of an advanced language change from below. In this section, I will use this information and the theoretical framework sketched in the first section of this chapter to formulate the hypotheses that will be tested in Chapter 6. Specifically, section 3.1 will introduce the main hypothesis. In section 3.2, I will argue that the results obtained in earlier studies for the reference of the NP and the verb tense may reflect two general cognitive constraints on linguistic expression: markedness of coding and statistical preemption. Additionally, I will introduce a third potential cognitive constraint on *haber* pluralization, namely, structural priming, which has not been considered in earlier work. In section 3.3, drawing on Labov's (2001) Principles of Linguistic Change, I will make precise predictions about the patterns of social covariation *haber* pluralization may feature.

3.1 Main hypothesis

Within the theoretical setting presented in section 1 of this chapter and taking into consideration the points raised in the introduction to this section, this study will explore the following main hypothesis:

> In Caribbean Spanish, the pluralization of presentational *haber* corresponds to a slowly advancing language change from below: the pluralized presentational schema with *haber* (<**AdvP** *haber* **Subj**>) is replacing the singular presentational construction with this verb (<**AdvP** *haber* **Obj**>). The variants only differ with regard to the syntactic function of the NP (singular variant: object; pluralized variant: subject) and the social types associated with their relative frequencies.

Of course, this is a very abstract hypothesis, which on its own, does not allow for any predictions. However, through reference to three cognitive factors (markedness of coding, statistical preemption, and structural priming) and Labov's (2001) Principles of Linguistic Change, a list of more detailed extrapolations can be drawn up. Let us consider these cognitive and social factors from up close, beginning with markedness of coding.

3.2 Cognitive factors

3.2.1 Markedness of coding

The typological literature indicates that human-reference nouns are more prototypical subjects than nouns that refer to other types of entities (Croft, 2003: 130; Dixon, 1979: 85). Therefore, in the light of the main hypothesis, the effect of human-reference NPs observed in earlier investigations may reflect the preference for unmarked coding, that is, the broader tendency for a "notion approximating an archetypical conception [to be] coded linguistically by a category taking that conception as its prototype" (Langacker, 1991: 298). This is captured by hypothesis 1.[28]

> *Hypothesis 1, Markedness of coding:* A more prototypical subject will more likely be coded as a subject. Conversely, a more prototypical object will more likely be coded as an object. This will lead speakers to select the pluralized presentational *haber* construction more often with NP arguments that are more similar to prototypical subjects and the singular presentational *haber* construction with NP arguments that are more similar to prototypical objects.

3.2.2 Statistical preemption

In Spanish, *haber* has always been used in a variety of constructions. With most of these, the verb is completely grammaticalized, meaning that it acts as an aspectual auxiliary. This is most evident for the perfect-tense construction, exemplified in Table 3.2.[29]

Table 3.2: Some compound paradigms of the verb *cantar* 'to sing'

	Present perfect	*Pluperfect*	*Subjunctive perfect*
1st singular	he cantado	había cantado	haya cantado
2nd singular	has cantado	habías cantado	hayas cantado
3rd singular	ha cantado	había cantado	haya cantado
1st plural	hemos cantado	habíamos cantado	hayamos cantado
3rd plural	han cantado	habían cantado	hayan cantado

Haber also functions as an auxiliary in two modal constructions. The first construction, <*haber de* infinitive>, expresses deontic obligation (see example 9), epistemic necessity (see example 10), prospectivity (see example 11), and, in certain varieties, futurity (see example 12) (Bello, 1860: §316; Jorge-Morel, 1978: 130; Real Academia Española & Asociación de Academias de la Lengua Española, 2009: §28.6ñ-q).[30]

[28] The hypotheses will be operationalized in Chapter 6.

[29] In Latin American Spanish, the second-person plural forms of verbs are not used.

[30] Henceforth in this chapter, Real Academia Española & Asociación de Academias de la Lengua Española (2009) will be referred to as RAE & ASALE (2009).

(9) Si *ha de crearse* un tribunal para juzgar los crímenes de guerra, tiene que ser absolutamente independiente (Real Academia Española, 2008b-, Spoken, Cuba).
'If a war tribunal *has to be created* to judge the war crimes, it should be completely independent.'

(10) Participant: Ella ya ha dado clases allí.
Interviewer: ¿Cuántos años?
Participant: Yo creo que *han de ser* cuatro con este (Davies, 2002-, Spoken, Mexico).
Participant: 'She has already given classes over there.'
Interviewer: 'How many years?'
Participant: 'I think that *it must be* four with this one.'

(11) Interviewer: Usted contó el Times Square a las doce de la noche de un Año Nuevo.
Participant: Seguramente. El Times Square. ¡Uy! es algo tan fantástico, tan... que *no he de olvidar* nunca en mi vida (Davies, 2002-, Spoken, Santiago de Chile).
Interviewer: 'You told me about Times Square at midnight of a New Year's Eve.'
Participant: 'That's right. Times Square. Wow! It's something so fantastic, so… that *I'm* never *going to forget* in my lifetime.'

(12) Para completar la lista que a continuación *hemos de incluir*, aparece recientemente una "iglesia satánica" en nuestro pueblo (Internet, Puerto Rico, http://goo.gl/Kypn1Q).
'To complete the list that *we will include* below, recently, a "Satanist church" has appeared in our town.'

The second modal construction with *haber* combines the verb with the complementizer *que* and an infinitive, as can be seen in example (13). Like <*haber de* infinitive>, this construction also expresses deontic obligation. There is, however, one difference: with <*haber que* infinitive>, the verb is used impersonally (RAE & ASALE, 2009: §28.6s-§28.6v).

(13) O sea, era muy distinto a antes. Anteriormente, esta gente era: "*Hay que hacer* esto, *hay que desbaratar* la universidad." Ahora no (Real Academia Española, 2008b-, Spoken, Puerto Rico).
'That is, it was very different from before. Before, those people were like: "This *has to be done*, the university *has to be destroyed.*" Now, they're not.'

Until the fifteenth/sixteenth centuries, *haber* was also widely used as a possessive lexical verb, and even today it may still be used like this (Álvarez-Martínez, 1996: 180; Hernández-Díaz, 2006: 1064, Note 2; Real Academia Española, 2005: s.v. *haber*). However, this use does not appear to have much vitality, because RAE & ASALE (2009: §4.13b) do not include possessive *haber* in their general overview of *haber* constructions. Indeed, various historical investigations have shown that by the end of the sixteenth century, *tener* is already the preferred possessive verb in Spanish (Fontanella de Weinberg, 1987: 33; Garachana-Camarero, 1997: 222; Hernández-Díaz, 2006: 1064). Still, RAE & ASALE (2009: §4.13d) state that the substandard first-person plural form *habemos* occurs sporadically with an abstract direct object and a possessive meaning. Their example (see example 14) and Fontanella de Weinberg's (1987:107) data suggest that, in Modern Spanish, possessive *habemos* and possessive *haber* in general only survive in idioms, such as *haber menester* 'to need' or *no haber remedio* 'to be a lost cause'. This is also evident from the Academies' comment that the use of possessive *haber* has to be seen as an archaic stylistic figure (Real Academia Española, 2005: s.v. *haber;* RAE & ASALE 2009: §4.13e).

(14) ¡Los hombres no *habemos* remedio! (Eduardo Labarca, *Butamalón*. From RAE & ASALE, 2009: §4.13d).
'We men, *we're a lost cause*!'

In legal documents and in literature, archaic *haber* also occurs as a participle (see example 15) and in passive constructions (see example 16). In these cases, the verb expresses meanings such as 'to arrive at', 'to achieve', 'to obtain', or 'to catch' (Bello, 1860: 257; RAE & ASALE, 2009: §41.6e-§41.6h).

(15) Según las estadísticas, de cada tres matrimonios *habidos* en el país uno fracasa, con impacto consiguiente en el fruto de los mismos (Davies, 2002-, Internet, Cuba).
'According to the statistics, out of three marriages *achieved* in the country, one fails, with consequent impact on the result of these [marriages].'

(16) No *pudo ser habido* el reo (from Bello, 1860: 257).
'The accused *could* not *be caught*.'

Finally, until the eighteenth century, *haber* also occurred in a presentational construction specifically dedicated to introducing time spans into discourse (see example 17), which later became supplanted by a competing structure with *hacer* (Fontanella de Weinberg, 1992b: 38).

(17) Cinco años *ha* que vine de las provincias del Perú con provisiones del marqués y gobernador Don Francisco Pizarro (Real Academia Española, 2008a-, 16th century).
'*It's been* five years since I came from the Peruvian provinces with provisions of the Marquis and governor Don Francisco Pizarro.'

Since *haber* has always occurred in multiple constructions, in the light of the main hypothesis, the fact that the pluralization of presentational *haber* does not occur as frequently with all tenses may reflect the general cognitive constraint 'statistical preemption', alluded to in section 1.2. As was explained in that section, when a form presents high token frequency in one construction schema, but only occurs sporadically in other patterns, it is taken to be stored as part of a partially lexically filled instance of this argument-structure construction with a much stronger representation than both the independent form and the abstract pattern (Goldberg, 1995: 79; Langacker, 1987: 59-60, 1991: 48). As retrieving this sub-construction from long-term memory requires relatively less effort, it disfavors the use of an alternative expression based on a competing construction schema that shares the same pragmatic and semantic constraints (Goldberg, 2006a: 94, 2009: 102-103).

Therefore, if certain tense forms of *haber* occurred mainly in the <**AdvP** *haber* **Obj**> pattern before <**AdvP** *haber* **Subj**> emerged as a conventional alternative, upon actuation of the change, the pluralized variant would not have been used frequently with those tenses. In subsequent generations, repetition usually ensures that this distribution remains intact (Bybee, 2006: 715). This leads to hypotheses 2a-2b.

Hypothesis 2, Statistical preemption:

> *Hypothesis 2a:* If the third-person singular form of a particular tense of *haber* was frequently used outside of the singular construction before presentational *haber* became involved in community-wide agreement variation, this verb tense will favor the pluralized presentational *haber* construction.
>
> *Hypothesis 2b:* The other verb tenses will disfavor the pluralized presentational *haber* construction.

Since these hypotheses assume that the preempting effect of certain verb tenses is a function of the degree of entrenchment of a particular form in the singular construction, hypothesis 2c follows quite naturally.

> *Hypothesis 2c:* When the need to encode an aspectually or modally more complex conceptualization forces speakers to construct a new expression involving aspectual or modal auxiliaries rather than retrieving a partially prefabricated expression from long-term memory, the tenses that were predominantly used in the singular presentational *haber* construction

before presentational *haber* became involved in community-wide agreement variation will favor the pluralized presentational *haber* construction.[31]

3.2.3 Structural priming

Socio- and psycholinguistic research has shown that once speakers have used or processed a constructional pattern, they tend to recycle it in the following stretches of discourse (Bock, Dell, Chang, & Onishi, 2007; Goldberg, 2006a: 120-125; Labov, 1994: 577-578; Pickering & Ferreira, 2008; Weiner & Labov, 1983). In the psycholinguistic literature, this is known as 'structural priming' (Pickering & Ferreira, 2008: 427). Therefore, if the variation amounts to a competition between two argument-structure constructions, as the main hypothesis claims, I expect to find the pattern described by hypothesis 3.

Hypothesis 3, Structural priming: The earlier mention of one of the presentational *haber* constructions in discourse will promote the use of the same construction in the next occurrence, regardless of variations in tense, aspect, or mood.

3.3 Principles of Linguistic Change

Against the background of the literature on *haber* pluralization, the main hypothesis claims that the alternation constitutes a slowly progressing language change from below. This claim entails the prediction that the alternations will display patterns of social and stylistic covariation typical of this type of linguistic evolution. A first such pattern is the 'apparent-time' distribution characteristic of linguistic changes (Labov, 1994:43-72),[32] which predicts the situation described by hypothesis 4.

Hypothesis 4, Apparent time: The youngest speakers will favor the pluralized presentational *haber* construction, whereas older speakers will make more use of the singular presentational *haber* construction.

However, the research reported in Chapter 2.2 suggests that *haber* pluralization may progress too slowly to be observed in apparent time (Díaz-Campos, 2003; Fontanella de Weinberg, 1992b). Therefore, more and less direct evidence may be necessary to test the change-in-progress hypothesis. In this regard, we may resort to Labov's (2001: Chap. 8) Gender Principle, which establishes that "[i]n linguistic change from below, women use higher frequencies of innovative forms than men do (Labov, 2001: 292). This leads to hypothesis 5.

---

[31] This hypothesis does not imply that frequent combinations of (singular) presentational *haber* and aspectual or modal auxiliaries cannot be stored as a single unit. Rather, it is inspired by the fact that aspectual/modal auxiliaries do not co-occur frequently with presentational *haber*. For example, in the twentieth-century section of *Corpus del Español* (Davies, 2002), there are only 232 presentational cases of third-person singular *poder haber* 'there can be' against 39,472 cases of third-person singular synthetic presentational *haber*. A similar pattern is found for *deber haber* 'there has to be', with 160 presentational third-person singular cases. Aspectual auxiliary constructions such as *acabar de haber* 'stop to be', *dejar de haber* 'stop to be', and *empezar a haber* 'start to be' do not provide any results.

[32] In sociolinguistics, the term 'apparent time' is used to refer to differences between generational groups in the use of variable linguistic phenomena (see Chapter 4.1.2.1).

> *Hypothesis 5, Gender Principle:* In comparison to men of the same social characteristics, women will use the pluralized presentational *haber* construction more often.

Yet, since gender-differentiated behavior is also found for changes from above (Labov, 2001: 274) and because the possibility of age-graded behavior[33] always exists for apparent-time distributions, more evidence will be needed before we can confidently conclude that this alternation constitutes a linguistic change from below. In this regard, the most conclusive indication of an ongoing change from below seems to be the social class distribution described by the Curvilinear Principle: "[l]inguistic change from below originates in a central social group, located in the interior of the socioeconomic hierarchy" (Labov, 2001: 188). This entails hypothesis 6.

> *Hypothesis 6, Curvilinear Principle:* The middle class will show higher frequencies of use of the pluralized presentational *haber* construction than the groups of lower and higher social status.

In addition, in changes from below, the innovative variants usually display no style shifting or increase in frequency when formality rises (Labov, 1972: 239, 2001: Chap. 3; Silva-Corvalán, 2001: 248-249). This leads to hypothesis 7.

> *Hypothesis 7, Formality:* When formality increases, the frequency of the pluralized presentational *haber* construction will not decrease.

Furthermore, Labov (1972: 138) observes that highly educated speakers, regardless of their social class, tend to conform to supralocal prestige norms, which suggests hypothesis 8.

> *Hypothesis 8, Educational achievement:* Higher educational achievement will favor the singular presentational *haber* construction, whereas a shorter formal education will promote the pluralized presentational *haber* construction.

Finally, if the variation is to be considered a change in progress, chances are high that the three speech communities under investigation will show different stages of the change, because

> [w]hen two speech communities are separated so that communication between them is reduced, then divergence is expected, and any degree of convergence requires an explanation (Labov, 2010: 5).

However, as hypotheses 1-3 propose that *haber* pluralization is constrained by the same set of general cognitive factors, some convergence is to be expected. In contrast, the social distribution of pluralized and singular presentational *haber* will probably be completely community-dependent. Hypotheses 9a and 9b attempt to capture this.

---

[33] Labov (1994) describes age grading as "a regular change of linguistic behavior with age that repeats in each generation" (Labov, 1994: 45).

Hypothesis 9, Divergence:

> *Hypothesis 9a:* The three Caribbean varieties will display different stages of the same evolution towards the pluralized presentational *haber* construction.
>
> *Hypothesis 9b:* The data will display similar tendencies for the cognitive factors, but the associations of the presentational *haber* constructions to social types will vary more significantly according to the respective speech community.

Before turning to the methods that were used in testing these hypotheses, let us first summarize the most important ideas that were presented in this chapter.

4. Summary

In this chapter, a concise sketch of Cognitive Construction Grammar has been presented. In this framework, speakers' grasp of their native language is pictured as a structured inventory of form-function pairings, called 'constructions'. The meanings of constructions refer to conceptualizations of things (nouns), events (verbs), qualities (adjectives and adverbs), or abstractions over events of the same type (argument-structure constructions), that is, ICMs. In syntax, argument-structure constructions provide psychologically plausible solutions for idioms and pragmatically or semantically motivated alternations. Additionally, constructions can be paired with social information, for which Cognitive Construction Grammar can model the social distribution of alternating constructions that refer to the same ICM and the way speakers use these distributions to position themselves against the background of social types and to express related social meanings.

Subsequently, the research questions and the hypotheses were introduced. Crucially, the main hypothesis of this study contends that *haber* pluralization amounts to a competition between two variants of the presentational construction with *haber*, which only differ in terms of their associations to social types and the syntactic function of their NP arguments. Assuming this main hypothesis and taking into account the trends observed in earlier investigations of *haber* pluralization, section 3 identifies four potential constraints on the variation: markedness of coding, statistical preemption, structural priming, and Labov's (2001) Principles of Linguistic Change. In Chapter 6, these hypotheses will be put to the test. However, first, the methods of this investigation should be introduced. These will be the topic of the next chapter.

4

Comparative Sociolinguistics

In recent years, the usage-based approach to language (e.g., Langacker, 1990: Chap. 10) has implied that cognitive linguists have increasingly moved away from introspective methods, in favor of corpus investigation, experimentation, and theorizing based on testable predictions (e.g., Geeraerts, 2006: 12; Gibbs, 2009; Grondelaers, Speelman & Geeraerts, 2009: 149-150; Heylen, Tummers, & Geeraerts, 2008: 91-92; Pütz, Robinson, & Reif, 2012: 244). Still, with a few exceptions (e.g., Clark, 2007, 2008; Hollmann & Siewierska, 2011; Kristiansen, 2008), spoken language variation and change largely remain under the radar of cognitive linguistics (Kristiansen, 2008: 47). This investigation intends to remedy this, adopting the comparative sociolinguistic methodology. Specifically, section 1 presents the decisions that were taken in sampling speakers from the three speech communities. Subsequently, section 2 discusses the fieldwork methods. Section 3 focuses on the transcription procedure, the selection of instances of *haber* and the 'envelope of variation'. Section 4, in turn, is dedicated to the statistical toolkit that will be used in Chapter 6. Section 5 introduces the comparative sociolinguistic method and the chapter concludes with a brief summary in section 6.

1. Judgment sample, selection criteria, and (post-)stratification

Corpora inevitably constitute limited samples of both the endless expressive possibilities a language has to offer and the usage patterns of all of its speakers. Therefore, it is important to define the samples that were considered for analysis sharply, as they will determine the robustness of the results to a large extent. To this end, section 1.1 introduces the 'judgment sampling' technique. Subsequently, section 1.2 will present the general criteria and the social characteristics according to which speakers were selected. Finally, section 1.3 focuses on the way social class is approached in this study and the procedure that was followed in post-stratifying the data by this parameter.

1.1 Judgment sample

Following standard practice in current variationist methodology (e.g., Milroy & Gordon, 2003: 30-33; Tagliamonte, 2006: 23-24), I sampled speakers from the Havana, Santo Domingo, and San Juan speech communities according to a number of previously set social characteristics. This is called 'judgment sampling' and the social characteristics used in this process are called the 'stratification variables'. With this method, the tacit linguistic norms of a speech community can be investigated successfully with a relatively small number of speakers. Of course, this requires that the analyst is realistic about the number and the types of hypotheses that can be explored with the data (Paolillo, 2013: 113-114), that the social categories are locally meaningful, and that the sample includes enough speakers to counterbalance speaker-specific idiosyncrasies (Almeida & Hernández-Campoy, 2005: 59; Johnson, 2009: 364; Labov, 2001: 326; Milroy & Gordon, 2003: 30; Tagliamonte, 2012: Chap. 4).

Concerning the latter issue, the literature suggests that three to five participants per cell created by the crossing of the stratification variables is sufficient (Almeida & Hernández-Campoy, 2005: 60; Chambers, 2009: 42-43; Feagin, 2002: 29; Milroy & Gordon, 2003: 30-35; Moreno-Fernández, 2003: 8; Tagliamonte, 2006: 23-24). In accordance with the guidelines of the *Proyecto para el Estudio Sociolingüístico del Español de España y de América* (PRESEEA) (Moreno-Fernández, 2003: 8), the samples of this study include three speakers per cell. Let us consider now the social characteristics that were used in selecting speakers.

1.2 Selection criteria and stratification variables

As a general requirement, in order to be eligible, all speakers had to be born and raised in their respective country and have lived in the capital for the last five years. Speakers meeting these requirements were then selected and grouped together according to their age (20-35 years vs. 55+ years), educational achievement (less than university vs. university), and gender (female vs. male). Additionally, as we will see in section 1.3, the samples were post-stratified by social class. However, before turning to that topic, in the remainder of this section, the local relevance of the stratification variables will be discussed, beginning with age.

1.2.1 Age

Age is a basic biological distinction that has some profound consequences for the roles the individual assumes in society and her/his expectations and views on life (Eckert, 1989: 246-247). In addition, age is essential to the study of linguistic change in progress, because contrasts between generational groups are generally assumed to reflect the historical development of the language. Of course, the 'apparent-time methodology' (as this construct is known in variationist sociolinguistics) rests on the

assumption that language remains more or less stable during adult life (Bailey, 2002: 320-324; Labov, 2001: 138).[1]

However, Chapter 2.2 suggests that the phenomenon advances at an extremely slow rate (D'Aquino-Ruiz, 2008; Fontanella de Weinberg, 1992b), if it progresses at all (Quintanilla-Aguilar, 2009). In this light, if any significant contrasts between generational groups will be found, these will probably only be visible between the youngest and the oldest age cohorts. Therefore, there is no need to sample all adult generations available. This way, one can also avoid including speakers aged one year older or younger in different generations, whereas there is no objective reason to assume that their speech is markedly dissimilar.[2] With this in mind, this study includes only two age groups, as is shown (1).

(1) • 20-35 years old
• 55 years and older

1.2.2 Educational achievement

Educational achievement, rather than social class was selected as the third stratification variable. This was motivated by both theoretical and practical concerns. Regarding the former, Milroy & Gordon (2003) question the usefulness of social class as a stratification variable in the Latin American context, "which is characterized by a large difference in access to power and advantage between the elite and the majority of the population" (Milroy & Gordon, 2003: 43). In this light, implementing a judgment sample with an equal representation of all social classes may render a severely disproportioned picture of the speech community. Additionally, since social class is usually defined as a function of multiple demographic parameters,[3] which may be valuated differently in the societies under study (Milroy & Gordon, 2003: 43),

[1] Earlier studies, which combined apparent-time investigation with a subsequent restudy of the same community after ten years or more, have proven this central assumption to be correct essentially (Chambers, 2009: 198-217; Labov, 1982: 67, 1994: Chap. 1). Additionally, considered on a whole, very few cases of age grading (i.e., a pattern of covariation with age that repeats itself generation after generation; Almeida & Hernández-Campoy, 2005: 40; Chambers, 2009: 207; Labov, 1994: 45) have been uncovered. Moreover, almost all involve phenomena typical of adolescent speech. Therefore, they are of less interest to apparent-time studies, which usually only include the adult population (Bailey, 2002: 324; Chambers, 2002: 358, 2009: 200; Milroy & Gordon, 2003: 36). Additionally, the claim that speakers' grasp of language remains more or less stable during adult life does not mean that usage is incapable of adding small, continuous readjustments to it, as is assumed in usage-based linguistics (Bybee, 2009: 347; Bybee & Beckner, 2010: 852). Rather, real-time replications of apparent-time investigations (e.g., Labov, 1994: Chap. 4) and longitudinal studies of individuals (e.g., Sankoff, 2004: 136) typically show some movement by adults in the direction of changes in progress. Yet, these studies also show that, although adults participate in ongoing changes, they do so "more sporadically and at a much lower rate than children" (Labov, 2010: 347). This suggests that once initial acquisition and stabilization of the vernacular has completed, the system will not be turned completely upside-down by changes in progress.

[2] For a similar approach, see Serrano (2006: 39). Of course, one can also avoid this by leaving, for example, a ten-year interval between the age groups, like, for example, Britain (1992) has done.

[3] For example, academic achievement, housing, income, etc. See section 1.3.

comparing samples stratified by social class may actually imply comparing, for all factors that are examined, the behavior of individuals whose only common feature is the social class index the investigator has superposed on the reality. This seems undesirable. Turning now to the practical issues related to the use of social class as a stratification parameter, it soon became evident while preparing the fieldwork that a multifactorial social class index would impose more restrictive selection criteria and, as a result, complicate the data collection process. Taking into account these three points, the sample was stratified by educational achievement.

When examining the influence of this parameter, Labov (1972: 138, 2010: 84) observes that university education triggers an important change in speakers' sensitivity to linguistic variation and its association to social types. Particularly for *haber* pluralization, Freites-Barros (2003: 380, 2004: 41) has shown that this increased sensitivity to linguistic variation causes university-educated speakers to judge pluralized *haber* as incorrect more frequently.[4] Therefore, this study only distinguishes two educational achievement levels (see 2), defined as the most advanced degree the speaker has obtained.

(2) • Less than university
• University[5]

1.2.3 Gender

Although speakers were selected according to their biological sex, in accordance with standard practice in current variationist sociolinguistics (e.g., Cheshire, 2002; Eckert, 1989), the oppositions between men and women will be approached in terms of gender, that is, in terms of the social types of masculinity and femininity. As Epstein (2007) observes,

> the gender divide is not determined by biological forces. *No society or subgroup leaves social sorting to natural processes*. It is through social and cultural mechanisms and their impact on cognitive processes that social sorting by sex occurs and is kept in place by the exercise of force and the threat of force, by law, by persuasion, and embedded cultural schemas that are internalized by individuals in all societies (Epstein, 2007: 4, emphasis in the original).

As a result, men and women typically assume different roles in society and society expects different styles of behavior from males and females (Cheshire, 2002;

[4] See Chapter 2.1.

[5] I tried to include as few students as possible, since they cannot be rated satisfactorily using this criterion. Only three participants were still pursuing a degree at the time of the interview. The first, a Puerto Rican young woman, was nearly graduating from her bachelor's degree, for which she was included in the university-educated group. The second participant, also a Puerto Rican young woman, had only recently started an associate's degree at a community college, for which I included her in the 'less than university' education group. Finally, the third participant, a Cuban young male, was in his last semester of law school (a six-year program in Cuba) when I interviewed him. Therefore, I rated him as a university graduate as well, because in Puerto Rico or the Dominican Republic he would have earned an undergraduate university degree already.

Chambers, 2009: 116, 140; Eckert, 1989: 246-247). For example, research in sociology shows that

> [f]emales are more likely than males to express concern and responsibility for the wellbeing of others, less likely than males to accept materialism and competition, and more likely than males to indicate that finding purpose and meaning in life is extremely important (Beutel & Mooney-Marini, 1995: 446).

These differences in behavior standards, expectations, and experiences have been proven to play an important role in linguistic change (e.g., Cheshire, 2002; Labov, 2001: Chap. 8, 12). For these reasons, gender is a must-have stratification parameter in all studies of language variation and change (e.g., Labov, 2001: 84; Tagliamonte, 2006: 23). Table 4.1 summarizes the sample as it was implemented in the three cities.

Table 4.1: Composition of the sample

Educational achievement	*25-35 years*		*55 + years*		*Total*
	Male	*Female*	*Male*	*Female*	
Less than university degree	3	3	3	3	12
University degree	3	3	3	3	12
Total	6	6	6	6	24

1.3 Post-stratification: social class

The samples were later post-stratified by social class.[6] In accordance with standard practice in variationist sociolinguistics (e.g., Chambers, 2009: 48-50; Labov, 1966/2006: 132-139; López-Morales, 1983: Chap. 1; Milroy & Gordon, 2003: 42-43), I constructed a composite social class index. To this end, two additional demographic parameters were registered, namely, housing and occupation. Besides these two variables, the index also includes a measure of educational achievement. For this measure, speakers who did not finish high school were separated from those who did, because the former group has greater prestige and easier access to power than the latter. This way, the list of possible scores provided in (3) was established for the three factors making up the social class index.

[6] At first glance, this might seem somewhat surprising, since I have argued in the previous section that social class may be problematic for comparative sociolinguistic analyses and studies performed in the Latin American context. However, the issues signaled in the previous section only arise when social class is used for the primary stratification of the data.

(3) • Educational achievement:
    0. Less than high school
    1. High school
    2. University
  • Housing
    0. Apartment or house in poor condition
    1. Small apartment or house (up to two bedrooms) in good condition
    2. Large apartment or house (three or more bedrooms) in good condition
  • Occupation (from Moreno-Fernández, 2003:9): [7]
    0. Peddlers and street vendors, unskilled workers, domestic workers, performers of services that do not require skilling
    1. Shop-owners, secretaries and clerks, craftsmen, artisans, mechanics, salesmen in shops, collectors/conductors, technical assistants, police officers and security guards, soldiers, truck, car, or bus drivers…
    2. University-educated professionals, teachers, small-scale entrepreneurs and producers, technicians, supervisors…
    3. Liberal professionals, mid-level managers of the public and private sector, commissioned and non-commissioned military officers, medium-scale entrepreneurs and producers, university professors…
    4. Senior officials of the executive, legislative, and judiciary branch, senior military officers, large landowners, senior executives of public and private sectors…

Because these three variables may not make an equal contribution to individuals' social status and because they may have different weights in the three communities (e.g., Milroy & Gordon, 2003: 43), I incorporated impact factors in the social class index. In order to establish these, after the recording sessions, I handed the participants a questionnaire (see Appendix A) with the instruction to rank, on a scale from one to five, educational achievement, housing, and profession by their importance for social status. This provided the average impact factors displayed in Table 4.2.

Table 4.2: Relative importance to social status of educational achievement, housing, and occupation in Havana, Santo Domingo, and San Juan

Factor	*Havana*	*Santo Domingo*	*San Juan*
Educational achievement	4.57	4.31	3.93
Housing	2.86	2.56	2.43
Occupation	3.29	3.63	4.29

[7] Housewives, retired participants, and workless participants were given the occupation score of the last formal job they had held.

Subsequently, each speaker's score for the three variables was first multiplied by the average impact factor that had been established for that variable and these weighted scores were summed together. Then, the raw numbers were converted into percentages of the maximum possible score. Finally, speakers were binned together in three groups by their prestige index scores: those who scored 0-33% were considered members of the lower class, those who scored 34-66% were considered to belong to the middle-class, and those who scored 67-100% were considered members of the upper class. This way, the social-class post-stratifications depicted in Table 4.3-Table 4.5 were obtained.

Table 4.3: The Havana sample, post-stratified by social class

Social class	*25-35 years*		*55 + years*		*Total*
	Male	*Female*	*Male*	*Female*	
Upper class	2	2	2	2	8
Middle Class	2	2	2	2	8
Lower class	2	2	2	2	8
Total	6	6	6	6	24

Table 4.4: The Santo Domingo sample, post-stratified by social class

Social class	*25-35 years*		*55 + years*		*Total*
	Male	*Female*	*Male*	*Female*	
Upper class	3	3	0	3	9
Middle Class	0	2	3	1	6
Lower class	3	1	3	2	9
Total	6	6	6	6	24

Table 4.5: The San Juan sample, post-stratified by social class

Social class	*25-35 years*		*55 + years*		*Total*
	Male	*Female*	*Male*	*Female*	
Upper class	2	2	3	3	10
Middle Class	1	3	1	2	7
Lower class	3	1	2	1	7
Total	6	6	6	6	24

2. Fieldwork methods

The fieldwork was carried out in March-April (San Juan), April-May (Santo Domingo), and May-June (Havana) 2011. In the three cities, the author, a fluent second-language speaker of Puerto Rican Spanish, conducted all the interviews. Most participants were volunteers recruited with the help of local consultants, who introduced the author as a student of the local culture and language on a class assignment. However, as it turned out to be impossible to fill out the quota with just volunteers, certain participants were rewarded cash incentives. Specifically, in Havana, one speaker was rewarded three convertible pesos (1 CUC=1 USD) for his participation. In Santo Domingo, ten participants received a 200-peso incentive (1 RD$=0.02 USD). Finally, in San Juan, four participants received a ten-dollar compensation.

The interviews were recorded using the rear-facing built-in 120° microphone of a Samson Zoom H2 digital recorder, set to 24bit/96kHz WAV format, with low-cut filter enabled and the microphone Auto Gain Control set to AGC 2 (Speech). The majority of the speakers have been recorded for about 40-120 minutes. The shortest interview span was of 29 minutes, the longest interview lasted two hours and 25 minutes, and the average duration oscillates around 60 minutes. The total amount of speech data that was collected sums about 76 hours or, roughly, 700,000 orthographic words.

The data were gathered combining three methods: a sociolinguistic interview, a story-reading task and a questionnaire-reading task. The motivation for this combination of methods is twofold. First, using story reading and, especially using questionnaire reading, speakers can be confronted with more variable contexts and with structures that occur too infrequently to be studied in a corpus of sociolinguistic interviews (Wolfram, 1986: 10). Second, combining semi-directed interviews, in which virtually no attention is turned to language, with two tasks that explicitly focus all attention on the speaker's speech habits creates an opportunity to observe style shifting (Labov, 1966/2006: Chap. 4, 1972: 98).[8] To this end, the interview and the two elicitation tasks were coded as, respectively, semi-spontaneous and elicited speech. These two styles are roughly similar to, on the one hand, Labov's (1966/2006: Chap. 4, 1972: Chap. 3) style B (interview style) and, on the other, his styles C (reading style) and D (word list style). Let us consider now the three data gathering methods, starting with the sociolinguistic interview.

[8] Labov (1966/2006: Chap. 4, 1972: 99) argues that the difference between formal and informal styles consists in that, in formal styles, more attention is paid to speech. This implies that, if we focus more attention on speech, speakers will automatically adopt a more formal style.

2.1 Sociolinguistic interview

As with most sociolinguistic interviews,[9] the goal of this part of the recording sessions was to obtain 30 to 45 minutes of relaxed speech from the participants as well as the full range of their demographic data. Following standard practice in variationist sociolinguistics, the interview evolved around thematic question modules, designed to invite the participants to talk about a particular topic for as long as they wanted (see Appendix B). The questions were inspired by Tagliamonte's (2006: Appendix B) updated version of Labov's (1966/2006: Appendix A) original interview schedule, the interview format of the PRESEEA project (Moreno-Fernández, 2003: 12-15), and the list of questions used by Quintanilla-Aguilar (2009: Appendix F).

Additionally, in order to investigate comprehension-to-production priming effects,[10] a set of questions with presentational *haber* (see example 4) was included in the thematic modules. In these questions, pluralized and singular presentational *haber* were used randomly.

(4) Interviewer: ¿Este, y *habían* castigos por no llevar el uniforme?
Participant: Sí, *había* castigos, si, si ibas con ropa de calle (LH03M12/LH264-LH265).
Interviewer: 'Er, and *were there*$_{Plur}$ punishments for not wearing the uniform?'
Participant: 'Yes, *there were*$_{Sing}$ punishments, if, if you dressed casually.'

2.2 Story-reading task and questionnaire-reading task

After the interview, the participants were instructed to read out loud a two-page text in which 31 decision contexts with presentational *haber* and some distractor verbs had been inserted (20 trials, 11 fillers; see Appendix C). As shown in example (5), while reading and without preparation, the participants had to choose the variant that corresponded to their own idiom.

(5) En una pequeña aldea, *había/habían* un anciano padre y sus dos hijos. El mayor era trabajador y llenaba de alegría el corazón de su padre, mientras el más joven sólo le daba disgustos…
'In a small village, *there were*$_{Sing}$/*there were*$_{Plur}$ an old father and his two sons. The oldest worked hard and filled his father's heart with joy, whereas the youngest only irritated him…'[11]

[9] See, for example, Labov (1966/2006: 87, 1984: 32-33).

[10] See Chapter 3.3.2.3 and Chapter 6.2.3.

[11] For the first two interviews of the Puerto Rican dataset, the format was somewhat different. For example, the first line of the text read: *En una pequeña aldea* _____(haber*, pasado*) 'In a small town _____ (*there to be*,

Since only basic literacy could be assumed for all speakers, the story-reading task was deliberately kept very simple. Rather than confronting the participants with a newspaper article or another kind of text written with an adult audience in mind, the story-reading task was based on a text written for children of about seven years of age: *Juan Sin Miedo* 'John Without Fear'. Still, as is shown in Table 4.6-Table 4.8, there were eight participants, mainly older speakers without university education, who were unable to complete the reading tasks on their own. In these cases, depending on the amount of time that had already passed by, the interview was either concluded (one participant from San Juan) or the interviewer read the text to the interviewees, instructing them to identify the form that corresponded to their own idiom (the seven other participants).

Table 4.6: Number of participants from Havana who completed the story- and questionnaire-reading tasks with the help of the interviewer, by age, educational achievement, and gender

Educational achievement	*25-35 years*		*55 + years*		*Total*
	Male	*Female*	*Male*	*Female*	
Less than university degree	0	0	2	1	3
University degree	0	0	0	0	0
Total	0	0	2	1	3

Table 4.7: Number of participants from Santo Domingo who completed the story- and questionnaire-reading tasks with the help of the interviewer, by age, educational achievement, and gender

Educational achievement	*25-35 years*		*55 + years*		*Total*
	Male	*Female*	*Male*	*Female*	
Less than university degree	0	1	1	0	2
University degree	0	0	1	0	1
Total	0	1	2	0	3

Table 4.8: Number of participants from San Juan who completed the story- and questionnaire-reading tasks with the help of the interviewer or did not complete the tasks, by age, educational achievement, and gender

Educational achievement	*25-35 years*		*55 + years*		*Total*
	Male	*Female*	*Male*	*Female*	
Less than university degree	0	0	1	1	2
University degree	0	0	0	0	0
Total	0	0	1	1	2

past tense)'. However, speakers turned out to have had extreme difficulties using these grammatical terms to insert the intended verb form. Therefore, the story-reading task was quickly adapted to its present format.

When it comes to the contexts that were presented to the participants, Chapter 2.2 suggests that *haber* pluralization is primarily conditioned by the absence/presence of negation, the characteristics of the NP argument, and the verb tense. However, incorporating two tokens of all possible combinations of these factors in the reading task would result in too large a number of trial items. Therefore, the task only includes a selection of verb tenses, multiple types of NP arguments, and affirmative and negative clauses.

Table 4.9 shows that, although not all tenses of *haber* could be represented in the task, there is an almost equal representation of the tenses that have been shown to disfavor *haber* pluralization in earlier research (9 tokens)[12] and those that have been shown to favor *haber* pluralization (11 tokens).[13] For the NP argument, Chapter 2.2 suggests that animacy is a major constraint on *haber* pluralization. Therefore, animate-reference nouns (8 tokens), and inanimate-reference nouns (12 tokens) are almost equally represented in the text, as are affirmative (11 tokens) and negative sentences (9 tokens).

Table 4.9: Forms of presentational *haber* included in the story-reading task, by animacy and the absence/presence of negation

Tense form	*Animate*		*Inanimate*		*Total*
	With negation	*Without negation*	*With negation*	*Without negation*	
Imperfect (*había/habían*)	0	2	2	1	5
Morphological future (*habrá/habrán*)	0	0	0	1	1
Periphrastic future (*va a haber/van a haber*)	0	0	1	0	1
Present perfect (*ha habido/han habido*)	0	2	0	0	2
Present tense (*hay/hayn*)	2	0	4	1	7
Preterit (*hubo/hubieron*)	0	2	0	0	2
Subjunctive present (*haya/hayan*)	0	0	0	2	2
Total	2	6	7	5	20

For the questionnaire-reading task, the participants were given a questionnaire consisting of 45 items (32 trials, containing 41 tokens of presentational *haber*, 13 fillers; see Appendix D) preceded by a description that evoked the usage context for the interpretation of the trial sentence, as can be seen in example (6). Then, they were instructed to read out loud the descriptions and the trial sentences, while simultaneously filling in the gaps with the multiple-choice answer that corresponded to their usage. The participants who were unable to complete the tasks without the help

[12] That is, the present and the preterit tense.

[13] That is, all other tenses.

of the investigator were not handed the full questionnaire, but rather a random selection of three to four pages (minimally 18 trials and 10 fillers).

(6) *Después de algún proyecto para mejorar la calidad del agua de las presas del país, un científico comenta:*
Hace diez años, no_____ más de tres sapos en esta presa. Ahora, cuenta con veinte patos, tres garzas y miles de peces.
a) hubo b) hubieron
'*Following a project to improve the water quality of the country's basins, a scientist comments:*
Ten years ago, _____ more than three frogs in this basin. Now, it has twenty ducks, three cranes, and thousands of fish.
a) there weren't$_{Sing}$ b) there weren't$_{Plur}$'

As signaled above, the main purpose of the questionnaire-reading task was to confront speakers with linguistic contexts that occur too infrequently in unscripted spoken language. Chapter 2.2 suggests two such types. First, there are the cases in which *haber* is not accompanied by a full NP, but rather by a direct-object pronoun. In order to evaluate whether speakers establish verb agreement with these pronouns, six tokens of *haber* + plural object pronoun were included into the questionnaire. Second, aspectual and modal temporal auxiliary constructions and the subjunctive tenses also occur rather infrequently in unscripted spoken language. As is shown in Table 4.10, the questionnaire-reading task includes multiple tokens of them. As was the case with the story-reading task, multiple tenses, types of NP arguments, and both negative and affirmative sentences were included in the questionnaire, but an equal representation of all combinations between these factors proved unfeasible. In this regard, Table 4.10 shows that the present and preterit tense represent a quarter of the tokens. When it comes to the NP arguments, animate-reference NPs make up 21 of the 41 presentational *haber* tokens. 11 out of the 41 tokens involve negation.

Table 4.10: Forms of presentational *haber* included in the questionnaire-reading task, by animacy and the absence/presence of negation

Form	*Animate*		*Inanimate*		*Total*
	With negation	*Without negation*	*With negation*	*Without negation*	
Acaba/acaban de haber ‘there has just been’	0	0	0	1	1
Conditional (*habría/habrían*)	1	1	1	0	3
Debía/debían haber ‘there must be’	0	1	0	0	1
Empezó/empezaron a haber ‘there has begun to be’	0	2	0	1	3
Empieza/empiezan a haber ‘there begins to be’	0	0	0	1	1
Imperfect (*había/habían*)	1	2	2	4	9
Morphological future (*habrá/habrán*)	0	1	0	2	3
Present perfect (*ha habido/han habido*)	1	1	0	1	3
Present tense (*hay/hayn*)	0	2	1	2	5
Preterit (*hubo/hubieron*)	1	3	1	0	5
Pudo/pudieron haber ‘there could be’	0	0	1	0	1
Seguirá/seguirán habiendo ‘there will continue to be’	0	1	0	0	1
Subjunctive imperfect (*hubiera/hubieran*)	1	0	0	2	3
Subjunctive present (*haya/hayan*)	0	1	0	0	1
Subjunctive present perfect (*haya habido/hayan habido*)	0	1	0	0	1
Total	5	16	6	14	41

3. Transcription, selection of cases, and envelope of variation

This section focuses on the procedures that were followed while processing the data. Particularly, section 3.1 describes the way the recording sessions were transcribed. Subsequently, section 3.2 introduces the decisions that were taken while selecting and coding the cases of presentational *haber* + plural NP. Section 3.3 describes the forms that were considered for analysis. In variationist sociolinguistics, this is called the 'envelope of variation'.

3.1 Transcription

Once the fieldwork was completed, the 72 recording sessions were transcribed in their full length using Microsoft Word 2011 for Mac and VideoLan Media Player. During this phase, two potential difficulties for the correct transcription of the (pluralized) cases of *haber* + plural NP were identified. First, Caribbean Spanish features three main allophones for the nominal plural marker /-s/: the alveolar sibilant [-s], the laryngeal fricative [-h], and a zero allophone (López-Morales, 1983: Chap. 3, 1992: 77-100).[14] At first glance, the latter could be problematic, as it could lead us to incorrectly interpret plural nouns as singular. However, research into this matter has shown that, in the majority of the cases, nominal plurality is redundantly marked at multiple sites in the NP and that speakers draw on cultural, phonological, pragmatic, and semantic information to resolve the number of the NP (Labov, 1994: 556-561; López-Morales, 1983: 55-57, 1992: 91-93; Poplack, 1984: 222). For instance, in example (7), the NP *tantos cafés y bares* 'so many cafés and bars' features three possible sites to mark plurality with [-s]: *tantos, cafés,* and *bares*. Of these three, the latter marks plurality unequivocally even without [-s] by the addition of plural [-e] to the stem /bar/. The plurality of the nominal can also be inferred from the coordination of the nouns *cafés* and *bares* and from the meaning of the indefinite quantifier *tantos* 'so many'. When such disambiguating information is not available, plural [-s] is rarely realized as zero (Labov, 1994: 561; Poplack, 1984: 210).

(7) Son años y como aquí *hay* tantos cafés y bares y, tú sabes, uno ha estado noches y noches, y horas y horas, y conversando sobre temas, y temas y temas y… (SJ12M12/SJ1391).
'I've been around here for years and since *there are*$_{Sing}$ so many cafés and bars around here, and you know, one has been out here for nights and nights, and hours and hours, and talking about topics, and topics, and topics, and…'

Therefore, the phonetic variation of /-s/ does not seem to impose severe methodological challenges. Still, according to Labov (1994: 562-563, 589) about 5%

[14] Additionally, /-s/ can also be assimilated to the following consonant (López-Morales, 1983: 38, 1992: 77).

of zero plural forms will be interpreted as singular, because the context fails to provide the necessary clues for the plurality of the form. Table 4.11 shows that this would imply that, among the cases collected with the sociolinguistic interview method, there should be some 35 plural tokens that cannot be recognized as such. However, as these represent less than 1% of the total number of cases of *haber* + plural NP, they would only have added some insignificant noise to the statistical analysis.

Table 4.11: Estimated number of (unrecoverable) zero plural markings in the sociolinguistic interview sections of the Havana, Santo Domingo, and San Juan corpora

	Havana	*Santo Domingo*	*San Juan*
Overall rate of use of zero /-s/[a]	21%	85%	38.2%
Cases of *haber* + plural NP in the sociolinguistic interview	450	507	401
Number of cases with zero plural marking[b]	95	431	153
Number of cases with unrecoverable zero plural marking[c]	4.73	21.6	7.7
Number of unrecoverable plural markings / total number of cases of *haber* + plural NP[d]	0.23% (N=4.73/2093)	1.17% (N=21.5/1842)	0.47% (N=7.7/1655)

Notes: [a] The overall rates of use of zero /-s/ are taken from Terrell (1979: 601) for Havana, from Terrell (1982: 305) for Santo Domingo, and from López-Morales (1983: 39) for San Juan. [b,c,d] The numbers provided in the table are estimates. Labov (1994: 86) indicates that /-s/ variation is a stable variable. The estimates are based on the supposition that the rates of use of the zero allophone have not changed significantly over the past forty years.

Second, Caribbean Spanish features three allophones for the verbal plurality morpheme /-n/: the alveolar nasal [-n], the velar nasal [-ŋ], and a zero variant with backward nasalization of the preceding vocal (López-Morales, 1983: 106, 1992: 121). Of these three, the velar nasal is the most frequent realization in the varieties of Havana, Santo Domingo, and San Juan (López-Morales, 1983: 109-110, 1992: 123-125). Therefore, in the vast majority of the cases, a clearly audible contrast exists between the absence and presence of /-n/, especially for verbs, for which plural /-n/ is almost never realized as zero (Poplack, 1984: 222).

However, for tokens followed by a nasal consonant, it proved difficult to differentiate the zero allomorph from cases of nasalization caused by assimilation with the ensuing consonant. In order to transcribe these cases correctly, I first slowed down the playback of the sound file to 10% of the original speed. Then, I compared the participant's pronunciation of the target form followed by a nasal consonant with her/his pronunciation of a zero plural followed by a non-nasal consonant. This showed

that, in the latter case, the vocal is already markedly nasal from the onset, whereas, in the former, it only becomes nasalized towards the onset of the consonant. This, in turn, helped identifying the absence or presence of /-n/. However, in spite of these efforts, the data might still display a margin of error similar to the one estimated for /-s/.

Finally, in order to make sure that all tokens of presentational *haber* had been transcribed correctly, I checked all the transcriptions against the sound files. Whenever disagreement emerged between the forms I heard the first and the second time, I marked the timing of the token. After transcribing all the interviews, these tokens were checked once more.

3.2 Selection of cases

While searching for tokens in the transcription files, it became evident that speakers hesitate frequently while completing the story- and questionnaire-reading tasks. This leads them to provide multiple contradictory responses to the same item, as is shown in example (8).

(8) Qué raro, esta mañana no, no *había*, *habían* más carros que otros domingos (SJ01M22/SJ161-SJ162).
'How strange, this morning *there weren't*$_{Sing}$, *there weren't*$_{Plur}$ more cars than on other Sundays.'

Therefore, a selection principle was established: only the speakers' final answers were taken into account for the quantification.[15] However, when the speaker repeated the same variant multiple times, all the tokens of that particular variant were quantified. For instance, for example (9), two tokens of *habrán* 'there will be$_{Plur}$' were coded for analysis.

(9) No es tu culpa tuya, es que siempre *habrán* unas per, *habrá*, *habrá* personas, *habrán* unas personas malas (SD02H21/RD275-RD278).
'It's not your fault, it's that *there will always be*$_{Plur}$ some pers, *there will be*$_{Sing}$, *there will be*$_{Sing}$, *there will be*$_{Plur}$ some bad people.'

Let us now turn to the contexts that are considered variable in this study, which will be the topic of the next section.

[15] In this case, *habían.*

3.3 Envelope of variation

In general, all contexts in which third-person pluralized or singular presentational *haber* is followed by a plural NP, including coordinated singular nouns, are considered variable. This includes the cases of *haber* followed by the noun *gente* 'people' when it is used as a plural count noun, as in example (10), meaning 'persons, individuals' (Real Academia Española, 2005: s.v. *gente*).

(10) *Habrán* gentes que lo hagan (SD05H11/RD594).
'*There will be*$_{Plur}$ persons who do it.'

This also includes the present-tense forms *hay-hayn.* As we will see in Chapter 6.2.2, this is motivated by the fact that my corpus provides 53 tokens of the vernacular plural *hayn*, which had already been documented in earlier investigations of Antillean Spanish (Holmquist, 2008: 28; Vaquero, 1996: 64). Therefore, if we want to follow the important 'Principle of Accountability',[16] the alternation between *hay* and *hayn* cannot be excluded from the scope of this investigation.

In contrast, first-person plural *haber* (see example 11) and the agreement variation displayed by the modal construction <*haber que* infinitive> '<have to infinitive>' (see example 12) are considered to be outside of the envelope of variation, even though these have also been treated as instances of *haber* pluralization in some surveys (e.g., DeMello, 1991; Freites-Barros, 2008; Holmquist, 2008; Quintanilla-Aguilar, 2009).[17]

(11) Y *habíamos* bastantes, bastantes estudiantes en, e, los salones de clase (SJ03H22).
'And *we were* plenty, plenty of students in, er, the classrooms.'

(12) Estamos trabajando y *hay que hacer* unas chapitas, ¿no? Entonces, mientras más rápido era mejor, porque *habían que pasarlas* por varias etapas y eran cantidades (Real Academia Española, 2008b-, Spoken, Puerto Rico).
'We are working and *one has to make*$_{Sing}$ badges, right. Well, it was the faster the better, because *they had to pass*$_{Plur}$ them through multiple stages and they were many.'

This is motivated by the fact that these two constructions do not refer to exactly the same conceptualization as third-person pluralized and singular *haber*. For first-person plural *haber*, the difference is rather subtle. It consists in that first-person plural *haber*

[16] The Principle of Accountability states that all occurrences of the alternation have to be included in the analyses, as opposed to only those that confirm the hypothesis that is put to the test (Labov, 1972: 72, 1982: 30, 1994: 550).

[17] My corpus does not provide any example of *haber que* pluralization. The twentieth-century section of Davies (2002-) only includes 4 tokens of pluralized *haber que* against a total of 7,429 tokens of the construction. This suggests that the phenomenon is rather infrequent.

includes the speaker in the *presentatum* (see example 11), whereas this is not the case for third-person pluralized *haber* (see example 13). As a result, first-person plural *haber* does not alternate with third-person singular *haber,* but rather with first-person plural *ser* or *estar*.

(13) Y *habían* bastantes, bastantes estudiantes en, e, los salones de clase (constructed example).
'And *there were*$_{Plur}$ a lot, a lot of students in, er, the classrooms.'

Finally, in the case of <*haber que* infinitive>, the contrast with presentational *haber* is quite clear, because this construction does not encode the POINTING-OUT ICM, but rather expresses deontic obligation.[18]

4. Statistical toolkit

Once all the cases of *haber* + plural NP had been selected and coded for the relevant cognitive and social factors, two statistical tools were used in order to evaluate their effect: mixed-effects logistic regression and conditional inference trees. Let us consider, briefly, these two.

4.1 Mixed-effects logistic regression

Cedergren & Sankoff (1974) introduced logistic regression analysis into sociolinguistics. Ever since then, generalized linear regression models fitted by maximum-likelihood estimation have been the statistical toolkit of choice among variationist sociolinguists (e.g., Johnson, 2009; Sankoff & Labov, 1979; Tagliamonte, 2006: 133-134).[19] Although many researchers continue to rely on this type of regression analysis, another alternative has appeared more recently: generalized mixed-effects regression models fitted by maximum likelihood estimation.

The main difference between these two types of regression models resides in the way group-internal variation is approached. That is, generalized linear models rest on the premise that there is no group-internal variation between words or individual speakers instantiating the same linguistic or social factors, which are called 'fixed factors' in statistical lingo. Therefore, when a particular variant occurs more or less frequently depending on whether a particular fixed factor is present, the regression model automatically infers that this factor has a favorable or an unfavorable effect on the occurrence rate of that variant. However, this does not exclude the possibility that (nearly) all tokens involving that fixed factor come from a few frequently occurring

[18] See Chapter 3.3.2.2.

[19] Researchers commonly speak of 'VARBRUL' (the name of Sankoff's original Fortran implementation of the algorithm, itself an abbreviation of *Variable Rule Analysis*) or 'GoldVarb' analysis (the name of the later C++ implementation for Mac and Windows by Sankoff, Tagliamonte, and Smith).

words or a few individuals, as is often the case.[20] Since these might not be the best exemplars of the linguistic or social categories they instantiate, generalized linear models potentially overestimate the influence of fixed factors (Johnson, 2009: 364; Tagliamonte, 2012: 130, 137; Tagliamonte & Baayen, 2012: 142-146).

In contrast, a generalized mixed-effects regression model allows us to specify a factor for the individual speakers of the sample and the specific words that occur in expressions, which are called 'grouping factors', 'random effects', or 'random intercepts' (Johnson, 2009: 363). These allow us to model the possibility that some words or some speakers[21] might favor or disfavor a particular variant over and above (or under and below) the contextual or social factors they instantiate (Johnson, 2009: 365; Tagliamonte, 2012: 137). Specifically, when a particular variant occurs more or less often when a particular fixed factor is present, a generalized mixed-effects model will first evaluate whether this effect is stronger than the variation between the individual speakers or words of the same group. If this is the case, the model will report that the fixed factor has a favorable or an unfavorable effect on the occurrence rate of that variant. In the opposing case, the model will conclude that the fixed factor has no influence whatsoever (Johnson, 2009: 365; Tagliamonte, 2012: 137). This renders generalized mixed-effects models highly suitable for handling "sociolinguistic data, drawn always from the production of individuals, inevitably from less than ideally distributed datasets, and with innumerable cross-cutting social and linguistic factors" (Tagliamonte, 2012: 139).

With this in mind, I performed generalized mixed-effects regression analyses with Rbrul (Johnson, 2014).[22] For the three datasets, I included the individual speakers and the lemmas of the nouns that occur with *haber* in the models as random intercepts. However, as these were collinear,[23] I had to run parallel analyses for the grouping factors (e.g., Baayen, 2008: 294). In Chapter 6, I will only report as statistically

[20] For example, in a study of the elision of unstressed vocals in Mexican Spanish, Serrano (2006: 48-49) finds that, when frequently occurring words such as *pues* 'well, as, since' and *antes* 'before' are excluded, the rates of elision drop by 50%.

[21] The focus on a community of individuals actually implies a return to Labov's early work. In his hindsight comments to the second edition of his PhD dissertation, Labov (1966/2006: 157) writes: "[m]any aspects of the NYC study influenced linguists' later work, but one aspect did not. There are no people in most of the sociolinguistic studies that followed – just means, charts, and trends. Although I have campaigned to bring people back into the field of sociolinguistics, there has been only a limited response on this front" (see also Paolillo, 2013: 89). Mixed-effects modeling brings the individual back to the foreground, while still focusing on the group.

[22] See Johnson (2009) for a discussion of the main features of Rbrul.

[23] Some lexical items only occur once in the corpus, for which the value of the grouping factor 'lemma' is partially predictable from the value of the grouping factor 'speaker'. In statistical modeling, this is called 'collinearity'. As regression analyses presuppose that factor groups operate independently from one another (Cedergren & Sankoff, 1974: 339-339; Labov, 1972: 231-232), both grouping factors cannot be included in the same regression model (Baayen, 2008: 198).

significant those fixed effects that proved to condition the variation for all speakers and all lexical items. In the tables of that chapter, probabilities below .50 disfavor the pluralized variant of presentational *haber*, whereas probabilities above .50 favor this variant. In the notes to the tables, two goodness-of-fit measures will be provided: the deviances of the models and their Akaike's Information Criteria. The values of both these measures should be as low as possible.

4.2 Conditional inference trees

Although the unequally distributed datasets typically used in sociolinguistic research are "the epitome of the type of data that mixed models are designed to handle" (Tagliamonte, 2012: 139-141), mixed-effects regression models may become less accurate when the data are distributed too unevenly across factor groups and represent multiple interactions between factors (Baayen, 2014: 363-364). In this sense, it can be useful to combine mixed-effects regression with another statistical approach that rests upon completely different distributional assumptions (Baayen, 2014: 364; Tagliamonte & Baayen, 2012: 161). If we achieve similar results with both approaches, we can be more confident that they are not due to distributional biases. Additionally, although a mixed-effects model provides insight into the influence of individual constraints while taking all other factors and intergroup variation into account, it says little about the way these constraints jointly determine speakers' choice for a particular variant (Tagliamonte & Baayen, 2012: 163).

These two concerns can be addressed at the same time with conditional inference trees (Baayen, 2014: 364; Tagliamonte & Baayen, 2012: 161, 164), which can be generated in R (R Core Team, 2013) with the *ctree()* function of the package *party* (Hothorn, Hornik, Strobl, & Zeileis, 2014).[24] According to Baayen,

> [c]onditional inference trees estimate a regression relationship by means of binary recursive partitioning. The *ctree* algorithm begins with testing the global null hypothesis of independence between any of the predictors and the response variable. The algorithm terminates if this hypothesis cannot be rejected. Otherwise, that predictor is selected that has the strongest association to the response, as measured by a *p*-value corresponding to a test for the partial null hypothesis of a single input variable and the response. A binary split in the selected input variable is carried out. These steps are recursively repeated until no further splits are supported (Baayen, 2014: 364).

In the conditional inference trees that will follow in Chapter 6.2.4, the ovals represent the factor groups. The higher a node is located in the tree, the stronger it is associated with speakers' choice between the presentational *haber* constructions. The branches that go down from the nodes represent the binary split the algorithm has established in the data. At the bottom, the thermometer graphs represent the proportion of pluralized

[24] See Tagliamonte (2012: 152-155) and Tagliamonte & Baayen (2012: Appendix A) for complete how-to guides.

presentational *haber* in light gray. Let us now consider the way these quantitative data will be compared in Chapter 6.2.5.

5. Comparative sociolinguistics

When describing the comparative sociolinguistic method, Tagliamonte (2002: 731, 2012: 166) indicates that researchers can draw on three lines of evidence to compare the use of variable linguistic phenomena in two or more varieties of the same language: statistical significance, relative strength of factors, and constraint ranking. If these three lines of evidence provide identical results for the varieties under comparison, this indicates that the same forces are at play in these varieties. The opposing case, in turn, will identify where the dialects are diverging and which constraints were originally found in their common ancestor, if any (Tagliamonte, 2002, 2006: 245-246).

The first line of evidence is perhaps the most self-explanatory, as it simply consists in evaluating whether the regression analyses select the same list of constraints as statistically significant for the varieties under comparison. The second line of evidence, in turn, consists in comparing the effect of individual factors within factor groups[25] and the magnitude associated with this effect.[26] Finally, the third line of evidence consists in ranking the factor groups by the strength of their impact on the variation, expressed in terms of the range comprised between the highest and the lowest factor weight of the group. This sort of ordered list of factor groups is called a 'constraint ranking'.

Although this method has provided interesting results (e.g., Claes, 2011; Tagliamonte, 2002, 2012: Chap. 6), range-based constraint rankings might not be the best instruments to assess the relative strengths of constraints, for two reasons. First, the range reported for a factor group depends on the number of factors included in that group, with factor groups comprising more factors typically reaching higher ranges. Second, because ranges derive from factor weights, the range obtained for any given factor group depends on the distribution of the data across the individual factors of the group, which may be very different for the samples under comparison.

These two issues can be addressed simultaneously by assessing the relative importance of factor groups with a random forest model of the variation, as Tagliamonte & Baayen (2012: 158-165) have suggested recently. According to Baayen, this type of statistical models unite

[25] That is, does factor X of the factor group Y favor or disfavor a particular variant?

[26] That is, how far does the probability value for the factor X of the factor group Y diverge from 0.5?

> a large number of conditional inference trees, resulting in a (random) forest of conditional inference trees. Each tree in the forest is grown for a subset of the data generated by randomly sampling without replacement from observations and predictors. The predictions of the random forest are based on a voting scheme for the trees in the forest: each tree in the forest provides a prediction about the most likely class membership, and the class receiving the majority of the votes is selected as the most probable outcome (Baayen, 2014: 366).

In R, random forests can be grown with the function *cforest()* of the package *party*. Once a random forest model has been established, we can derive the relative impact of the different factor groups by calculating the loss in prediction accuracy of the model when the factors of a group are randomly permuted, breaking the associations between the dependent variable and the factors of the group. In R, this can be achieved with the function *varimp()* of the same package. The greater the loss in prediction accuracy, the more important a factor group is (Baayen, 2014: 366; Tagliamonte & Baayen, 2012: 160).[27] Let us now summarize the most important ideas that were presented in this chapter.

6. Summary

This chapter has outlined the methodological framework of this study. Most importantly, we have seen that this investigation draws on a judgment sample of three times 24 speakers, equally divided over two education groups, two gender groups, and two age groups. Furthermore, I have explained that the data were collected using a combination of semi-directed interviews with two elicitation tasks, which can be taken to represent semi-spontaneous and elicited speech. Additionally, the advantages of mixed-effects logistic regression using Rbrul have been highlighted and it was suggested that conditional inference trees constitute ideal companions for this type of regression analysis (Baayen, 2014; Tagliamonte & Baayen, 2012). Finally, I have introduced Tagliamonte's (2002, 2006: 245-247, 2012: Chap. 6) comparative sociolinguistic method, including Tagliamonte & Baayen's (2012) recent suggestions regarding the use of conditional variable permutation in random forests to establish constraint rankings. In Chapter 6, these statistical techniques will be used to elaborate a comprehensive analysis of *haber* pluralization in Caribbean Spanish. However, before turning to the statistical analysis of the corpus, in the following chapter, the pragmatic, semantic, and syntactic properties of the presentational *haber* constructions will be discussed.

[27] See Tagliamonte & Baayen (2012: Appendix A) or Tagliamonte (2012: 144-156) for step-by-step guides.

Part B

Results

This page intentionally left blank

5

A Cognitive Construction Grammar approach to presentational haber

In Chapter 3.3.1, the main hypothesis introduces the claim that the pluralization of presentational *haber* results from a competition within the grammar-lexicon between two variants of the presentational construction with this verb (<**AdvP** *haber* **Obj**> and <**AdvP** *haber* **Subj**>), which allows speakers to position themselves in terms of social types. However, the pragmatics, semantics, and syntax of the presentational *haber* constructions and their potential differences in these respects have so far largely remained without discussion. Therefore, this chapter will provide an overview of the characteristics of the constructions, which will allow me to identify any possible contrasts between them. Particularly, section 1 is concerned with the meaning of presentational *haber*. Subsequently, section 2 focuses on the nominal argument. Section 3 deals with the status of the adverbial phrase that appears frequently with presentational *haber*. Section 4, in turn, introduces the conditions that constrain the use of implicit nominal arguments and adverbial phrases. Finally, in section 5, a brief summary is presented.

1. The meaning of the presentational *haber* constructions: POINTING-OUT

The literature suggests that the meaning of presentational *haber,* like that of all presentational constructions, refers to a cognitive routine that introduces a nominal entity into discourse, asserting its existence, and situating it in a mental space[1] (Bolinger, 1954: 334, 1977: 92-93; Hernández-Díaz, 2006: 1056; Lakoff, 1987: 554;

[1] Fauconnier defines mental spaces as "small conceptual packets constructed as we think and talk for the purpose of local understanding and action" (Fauconnier & Turner, 1996: 113), and as such, they belong to the realm of working memory (Fauconnier, 2007: 351). In other words, mental spaces are novel, temporal conceptualizations that organize the information speakers and hearers are presented with in usage events. This includes the base space, the common ground shared by the hearer and the speaker (Croft & Cruse, 2004: 33). New mental spaces are built up dynamically in working memory by mixing fragments of other mental spaces with procedural and factual knowledge (Fauconnier & Turner, 1996: 115, 2006: 309-310). This process is called 'blending' and the output spaces are called 'blends' (Fauconnier, 2007: 351-352).

Langacker, 1991: 352-353; Suñer, 1982: 95). This is captured in Lakoff's (1987), POINTING-OUT ICM:

> [i]t is assumed as a background that some entity exists and is present at some location in the speaker's visual field, that the speaker is directing his attention at it, and that the hearer is interested in its whereabouts but does not have his attention focused on it, and may not even know that it is present. The speaker then directs the hearer's attention to the location of the entity (perhaps accompanied by a pointing gesture) and brings it to the hearer's attention that the entity is at the specified location (Lakoff, 1987: 490).

Chapter 2.2 has shown that pluralized and singular presentational *haber* are interchangeable in every context. Therefore, the main hypothesis claims that the two constructions encode the same ICM,[2] which implies that they are subject to the same pragmatic and semantic constraints. The remainder of this chapter will try to establish whether these two claims are justified.

2. The nominal argument

In this section, the characteristics of the nominal arguments of pluralized and singular presentational *haber* will be examined in the light of examples drawn from the corpus, Davies (2002-), the Internet, and Real Academia Española (2008b-). Particularly, section 2.1 will be concerned with its argument role, information status, and semantic function. Then, in section 2.2, its syntactic properties will be investigated.

2.1 Argument role, information status, and semantic function

Because the POINTING-OUT ICM only describes the act of bringing a referent "out of limbo into presence" (Bolinger, 1954: 335), the nominal encodes virtually the entire conceptual import of the clause. Semantically, we can conceive of this element as the trajector of the clause, which, however, is merely present in the scene that is being presented through the construction. Therefore, it is probably safe to assume that it is assigned a 'zero' argument role (Langacker, 1991: 288). Examples such as (1) and (2) show that this is the case for both pluralized and singular presentational *haber*.

(1) Después que mataron a Trujillo, pues fue Trujillo que trajo esa gente. Y trajo españoles también. *Habían* colonias españolas aquí (SD16H22/RD2200).
'After they killed Trujillo, because it was Trujillo who brought those people here. And he also brought Spaniards. *There were*$_{Plur}$ Spanish colonies here.'

(2) Pero que sí que *hubo* muchas, muchas casas, e, destrozadas, muchas casas, e, desaparecidas (SD23H12/RD3065).
'But that *there were*$_{Sing}$ many, many, er, destroyed houses, many, er, disappeared houses.'

[2] In other words, do the same referential work.

In turn, from the POINTING-OUT ICM it follows that, in affirmative expressions, the NP of presentational *haber* can only be interpreted as referring to a specific referent (Prince, 1992: 299-300) unless, as we will see below, the *presentatum* is explicitly construed as a type. Again, examples such as (3) and (4) suggest that this is the case for both variants of the presentational *haber* construction.

(3) En Salcedo *habían* muchos árabes, que le decían 'turcos', porque Turquía, e, parece que dominaba los países árabes y tenían mucha represión (SD16H22/RD2210).
'In Salcedo, *there were*$_{Plur}$ many Arabs, who were called 'Turks', because it appears that Turkey, er, dominated the Arab countries and they had much repression.'

(4) Este, vivienda acá *hay* muchos condominios (SJ02M12/SJ168).
'Er, housing, here, *there are*$_{Sing}$ many condominiums.'

Regarding information status, POINTING-OUT places stringent constraints on the nominal argument, as this ICM implies that the NP of the presentational *haber* constructions cannot be 'given'[3] with respect to the hearer's beliefs, consciousness, or world knowledge. Indeed, pluralized and singular presentational *haber* co-occur most often with indefinite nominal arguments, as is shown in examples (5)-(7) (Bull, 1943: 122; Fernández-Soriano & Táboas-Baylín, 1999: 1755-1756; Real Academia Española & Asociación de Academias de la Lengua Española, 2009: §12.2l, §20.2g, §20.3f-h;[4] Utley, 1954).

(5) Ahora no. Ahora no *hay* principios (SD02H21/RD230).
'Not nowadays. Nowadays, *there aren't*$_{Sing}$ principles.'

(6) En aquella época *había* cinco, seis millones de cubanos (LH12H21/LH1628).
'At that time, *there were*$_{Sing}$ five, six million Cubans.'

(7) Y desde luego, e, cantidad, no sé qué decirte, pero me imagino que sobre todo en el campo, *deben haber* muchos más, muchas situaciones de esa naturaleza (SJ03H22/SJ336).
'And, of course, er, quantity, I don't know what to tell you, but I imagine that, mostly on the countryside, *there must be*$_{Plur}$ many more, many situations of that nature.'

However, cases such as example (8) and example (9) illustrate that this does not mean that the presentational *haber* constructions only allow indefinite, discourse-new NPs, as

[3] In the sense of Chafe (1974, 1976).

[4] Henceforth in this chapter Real Academia Española & Asociación de Academias de la Lengua Española (2009) will be referred to as RAE & ASALE (2009).

some authors have argued (e.g., Fernández-Soriano, 1999: 131; Fernández-Soriano & Táboas-Baylín, 1999: 1755; Freeze, 1992: 557).

(8) Y *habían* los almuerzos, iban los tíos míos, iban los primos (SJ14H22/SJ1681).
'And *there were*$_{Plur}$ the lunches, my uncles went, my cousins went.'

(9) Bueno, sí, aquí en Cuba *hay* todas esas cosas (LH21H11/ LH2866).
'Well, yes, here in Cuba, *there are*$_{Sing}$ all these things.'

Rather, pluralized and singular presentational *haber*, like English presentational *there is/there are* (Lakoff, 1987: 545; Prince, 1992: 301; Ward & Birner, 1995: 740), seem to allow definite/discourse-old NPs,[5] provided they refer to entities that are (or can be construed as) new to the hearer (González-Calvo, 2002: 649-650; Suñer, 1982: 97-100).

In particular, when discussing the co-occurrence of English presentational *there is/there are* with definite NPs, Ward & Birner (1995: 730) identify five types of definite nominal arguments that can occur with this construction. As the list in (10) shows, many of these are also hearer-old.

(10) 1. Hearer-old entities treated as hearer-new
2. Hearer-new tokens of hearer-old types
3. Hearer-old entities newly instantiating a variable
4. Hearer-new entities with uniquely identifying descriptions
5. False definites[6]

Although Ward & Birner (1995) base their conclusions on English presentational *there is/there are,* the fact that the presentational *haber* constructions fulfill the same discourse function suggests that they might represent a similar behavior. Additionally, because "[d]ifferences in the packaging of information are perhaps the most important reason why languages have alternative ways to say 'the same' thing" (Goldberg, 2006a: 129-130), we might find some contrasts here between the two presentational *haber* constructions. Let us consider this matter from up close.

2.1.1 Hearer-old entities treated as hearer-new

The first type of hearer-old entities that can occur with English presentational *there is/there are* according to Ward & Birner (1995) are the hearer-old entities that are treated as hearer-new. In the literature, these are usually labeled 'reminders'. As the label of this type of definites implies, their felicitous usage requires that the speaker

[5] In the sense given to these terms by Prince (1992).

[6] Ward & Birner (1995) use the term 'false definites' to refer to syntactically definite NPs that encode discourse-new/hearer-new information.

assumes that the hearer has forgotten, at least temporally, about a referent[7] that has already been evoked in earlier discourse (Bolinger, 1977: 115-117; Lakoff, 1987: 545, 561; Ward & Birner, 1995: 750). In other words, with reminders, the use of a presentational construction is licensed by the fact that the speaker assumes that the hearer has forgotten about the referent of the NP, whereas the use of the definite determiner is motivated by the speaker's expectation that the hearer can at least recognize the entity that is being reintroduced (Suñer, 1982: 85; Ward & Birner, 1995: 730-731). For this reason, in English, it is most common to form reminding expressions with demonstratives, as in example (11), rather than with definite articles, which would imply that the hearer is expected to recall the referent (Langacker, 1991: 98; Ward & Birner, 1995: 731).

(11) She is running as an insider. That is a mistake. Then there are just the stray gaffes. She said, in a famous episode, she was asked to go shake hands. And she said: "Well, actually, I. No. What would I gain by shaking hands out of Fenway Park?" Well, that is exactly what you should be doing. You're running for office. Shake some hands. Go out and meet some people. So*, there is* that problem (Davies, 2008-, Press).

In Spanish, reminders are usually constructed with *tales* 'such', as in examples (12) and (13), or the close-to-hearer demonstrative *esas/esos* 'these', as in examples (14) and (15) (Suñer, 1982: 85-86). The examples also suggest that reminding definite NPs are more common in expressions involving negation.

(12) Entonces, esto fue creando en la gente un conocimiento de que no *habían* tales milagros (Real Academia Española, 2008b-, Fiction, Colombia).
'Well, this began to create in the people a knowledge that *there weren't*$_{Plur}$ such miracles.'

(13) Nadie tiene derecho a estar provocando a otro país, enviando aviones con el pretexto de rescate. *No había* tales pretextos de rescate, lo que había era la promoción de las salidas ilegales del país (Real Academia Española, 2008b-, Spoken, Cuba).
'No one has the right to be provoking another country, sending aircraft under the pretext of rescue operations. *There weren't*$_{Sing}$ such rescue pretexts, what there was, was the promotion of illegal exits out of the country.'

[7] It should be observed that it is the referent that is reintroduced to the hearer, not the specific noun, such that, for example, *these boys* could be used in a reminding expression to reintroduce, say, *little Johnny and Bert*.

(14) Interviewer: Y ese nuevo orden enriquece la vida del hogar, fíjate. Es más televisión, más radio, más cosas dentro de la casa y enriquece la vida del hogar y por lo tanto, pues, debilita la otra.
Participant: Debilita la otra, sí. Antes no *habían* esas cosas y tenía que el individuo irse a la calle a, a procurarse la diversión (Davies, 2002-, Spoken, Puerto Rico).
Interviewer: 'And this new order enriches the life of the home, notice that. It is more television, more radio, more things inside the house and this enriches the life of the home, and, therefore, debilitates the other.'
Participant: 'It debilitates the other, yes. Before *there weren't*$_{Plur}$ these things, and the individual had to go out on the streets to, to get diversion.'

(15) Interviewer: En los libros sobre el español de Santo Domingo se dice que de vez en cuando en el interior del país todavía se usa 'su merced' o, 'vuestra merced'. ¿Usted alguna vez lo ha escuchado?
Participant: No.
Interviewer: ¿No?
Participant: No, eso no es verdad.
…
Participant: Yo no he oído esto, a, aquí no *hay*, aquí no *hay* ese, esos términos: 'su merced', 'vuestra'. No, eso no es verdad (SD15M21/ RD1908- RD1909).
Interviewer: 'In books about the Spanish of Santo Domingo, they say that sometimes in the interior of the country they still use 'your grace'. Have you ever heard it?'
Participant: 'No.'
Interviewer: 'No?'
Participant: 'No, that's not true.'
…
Participant: 'I haven't heard this, he, here, *there aren't*$_{Sing}$, here, *there aren't*$_{Sing}$ that, those terms: 'your grace', 'your'. No, that's not true.'

In my data, reminding expressions are also formed with the close-to-speaker demonstrative *estos* 'these', but only with the pluralized construction, as can be seen in example (16). However, when we extend the range of data that are considered, singular tokens can also be found, as is shown in example (17).

(16) Interviewer: ¿Este, cuando se mudó aquí, habían cosas a las que tuvo que acostumbrarse?
Participant: ¿Tales cómo?
Interviewer: No sé por ejemplo. E, este, no sé, cosas.
Participant: Este, bueno, no, me tuve que acostumbrarme a vivir en un condominio cuando yo viví en una casa.
…
Interviewer: ¿Y usted recuerda como la ciudad era antes? O sea, cuando era niña.
Participant: Cuando yo era niña, sí. *No habían* estos condominios, desde luego (SJ01M22/ SJ07).
Interviewer: 'Er, when you moved here, were there things you had to get used to?'
Participant: 'Such as?'
Interviewer: 'I don't know, for example. Er, er, I don't know, things.'
Participant: 'Er, well, no, I had to get used to living in a condominium, when I had always lived in a house.'
…
Interviewer: 'And do you remember what the city was like before? That is to say, when you were a girl.'
Participant: 'When I was a girl, yes. *There weren't*$_{Plur}$ these condominiums, of course.'

(17) Estudiaba el que tenía plata, el que no, no podía estudiar. Qué diferencia ¿no? Esos bochinches estudiantiles.
…
Yo entiendo que no es lo mismo. En ese tiempo Mérida contaba cuatro, cinco mil habitantes, era un pueblito. Y no *había* estos bochinches porque no había nadie, era el trabajo, todo el mundo pegado al trabajo y esas cosas, y nadie estaba pensando en hacerle mal al otro (Real Academia Española, 2008b-, Spoken, Venezuela).
'Those who had money studied, those without could not study. What a difference, right? These student riots.
…
In my understanding, it's not the same. At that time Merida had four, five thousand inhabitants, it was a small town. And *there weren't*$_{Sing}$ these riots because there was nobody, everyone was occupied with their jobs and these things, and no one was thinking about doing harm to another.'

2.1.2 Hearer-new tokens of a hearer-old type

Ward & Birner (1995: 732-733) observe that the definite NPs of English presentational *there is/there are* constructions can also be licensed if the NP introduces a hearer-new instance of a known or inferable type. As various authors point out, this reading requires an adjective that construes the NP in this way (Lakoff, 1987: 546; Ward & Birner, 1995: 732-733). With Spanish presentational *haber,* this function may be fulfilled by adjectives such as, among others, *mismas/mismos* 'same' (see examples 18 and 19), *necesarias/necesarios* 'necessary' (see examples 20 and 21), *obligatorias/obligatorios* 'obligatory', and *suficientes* 'sufficient' (see examples 22 and 23) (RAE & ASALE, 2009: §15.6l-ñ; Torrego-Salcedo, 1999: 1795).

(18) Ya en Azteca *habían* los mismos comentarios acerca de su manera prepotente y payasa (Internet, Message board, Mexico, http://goo.gl/9RWHLP).
'Already in Azteca,[8] *there were*$_{Plur}$ the same comments about her overbearing and clownish way.'

(19) En las paredes *había* los mismos mapas de acrílico transparente con las fronteras en negro (Internet, Magazine, Argintina, http://goo.gl/I8auJE).
'On the walls *there were*$_{Sing}$ the same maps of transparent acryl with the borders in black.'

(20) Por motivos de economía, en México nunca *han habido* los filtros necesarios para tamizar esos testimonios (Internet, Press, Mexico, http://goo.gl/Oh2Kp6).
'For economical reasons, in Mexico, *there have* never *been*$_{Plur}$ the necessary filters to sift these testimonies.'

(21) Se habló de que por lo menos unos 200 mil trabajadores en estas condiciones pasarían a las filas de la formalidad, cosa que no ha sucedido en gran parte por la baja en el crecimiento económico y porque no *ha habido* los incentivos necesarios (Internet, Press, Mexico, http://goo.gl/HDZDVy).
'There was talk that at least some 200 thousand workers in these conditions would pass to the ranks of formality,[9] something that has not happened, to a large extent because of the decrease in the economic growth and because *there have not been*$_{Sing}$ the necessary incentives.'

(22) Simplemente no *habían* los suficientes trabajadores estadounidenses para recoger las cosechas a precios que las hubiesen hecho rentables (Internet, Blog, Ecuador, http://goo.gl/HUfGvb).
'Simply *there weren't*$_{Plur}$ the necessary amount of American workers to collect the harvests at prices that would have made them profitable.'

[8] A Mexican soap opera production house.
[9] That is, would see their jobs officialized.

(23) Al parecer, ese día, ya *había* los suficientes equipos como para dar paso a la primera fecha del torneo (Internet, Press, El Salvador, http://goo.gl/jv9IIQ).
'As it appears, that day, *there were*$_{Sing}$ already the necessary amount of teams to proceed with the first date of the tournament.'

Anaphoric pronouns, as in example (24), are also interpreted as new instances of a referent type evoked earlier. Here, we would expect to find only the singular construction, because the pronouns that appear with presentational *haber* are accusatives. However, as we will see in Chapter 6.2.3, the corpus also provides a limited number of pluralized tokens involving direct-object pronouns. As I will show, rather than invalidating the main hypothesis, the fact that tokens such as example (25) typically occur after the interviewer or the participant have used a pluralized presentational *haber* expression appears to suggest that structural priming causes individual speakers to reanalyze the direct-object pronoun as a subject pronoun.

(24) Sí, sí, aquí también los *hay*. Y yo supongo que los *habrá* en, en Bélgica, en, en, en Italia, en todos lados (LH15H21/LH1596).
[Yes, yes, here, them$_{Acc}$ *there are*$_{Sing}$ as well. And I suppose that them$_{Acc}$ *there will be*$_{Sing}$ in, in Belgium, in, in, in Italy, everywhere.]
'Yes, yes, here, *there are*$_{Sing}$ as well. And I suppose that *there will be*$_{Sing}$ in, in Belgium, in, in, in Italy, everywhere.'

(25) ¿A, acá ya *habían* carros? ¡Claro que los, claro que los *habían*, no, no soy tan viejo!" (LH20H12/LH2765-LH2766).
[*W, were there*$_{Plur}$ already cars here? Of course that them$_{Acc}$, of course that them$_{Acc}$ *there were*$_{Plur}$, not, I'm not that old!]
'*W, were there*$_{Plur}$ already cars here? Of course that them$_{Acc}$, of course that *there were*$_{Plur}$, I'm not, I'm not that old!'

My corpus also provides pluralized and singular examples of presentational *haber* expressions introducing a hearer-new token of a hearer-old type with the generic possessive determiner *sus* 'one's', implying 'the/your typical', or 'the usual' (see examples 26 and 27).

(26) Interviewer: ¿Este, y *habían* platos que tu madre te hacía especialmente para ti porque te gustaban tanto?
Participant: T, sí, ha, *habían* sus boberías pero no mucho. Mi casa nunca fue una casa que tuvo grandes posibilidades (LH21H11/LH2842).
Interviewer: 'Er, and *were there*$_{Plur}$ dishes that your mother made especially for you because you liked them so much?'
Participant: [T, yes, the, *there were*$_{Plur}$ one's silly things, but not much. My home was never a home that had large possibilities.]
Participant: 'T, yes, the, *there were*$_{Plur}$ the/your typical silly things, but not much. My home was never a home that had large possibilities.'

(27) E, sí, siempre *había* sus diferencias y sus celos, pero nos criamos bien (SD23H12/RD3049).
[Er, yes, *there were*$_{Sing}$ always one's differences and envies, but we grew up alright.]
'Er, yes, *there were*$_{Sing}$ always the/your typical differences and envies, but we grew up alright.'

Additionally, since types are usually encoded with indefinite NPs, presenting hearer-new tokens of hearer-old types does not necessarily involve the use of definite determiners in my data. Rather, it is quite common in Spanish to use plural or singular presentational *haber* expressions with implicit NPs to present new instances of a previously evoked type, as is shown in examples (28) and (29).

(28) Interviewer: ¿En esa época se veía que los hermanos estaban corrigiendo a, a sus hermanas?
Participant: *Habían, habían, habían*, pero mi mamá decía que no era correcto eso (SD15M21/ RD1858- RD1860).
Interviewer: 'In that time, would you see that brothers were correcting their sisters?'
Participant: '*There were*$_{Plur}$*, there were*$_{Plur}$*, there were*$_{Plur}$, but my mom said that that was not correct.'

(29) Interviewer: ¿No? ¿Este, no, no había peleas en aquel entonces?
Participant: No, siempre *hay*, lo que pasa es que, o no me han tocado, o yo no he querido estar (LH08H12/LH988).
Interviewer: 'No? Er, no, weren't there fights back then?'
Participant: 'No, *there are*$_{Sing}$ always, what happens is that, either they have not affected me, or I didn't want to be involved.'

Finally, speakers may also specify the amount of new tokens of the type they wish to bring to the hearer's attention by using a quantifying pronoun or another quantifying expression, as is shown in examples (30) and (31).

(30) Interviewer: ¿Y habían cosas que no le gustaban de la ciudad?
Participant: *Habían* algunas que no me gustaban, sí, sí (LH17M21/LH2291).
Interviewer: 'And were there things that you didn't like about the city?'
Participant: '*There were*$_{Plur}$ a few that I didn't like, yes, yes.'

(31) Interviewer: ¿Este, entonces, que usted recuerde cuando usted era niño, habían más conversaciones en la calle cuando usted era niño?
Participant: Claro que sí, *había* más (SD02H21/RD14).
Interviewer: 'Er, well, for what you can remember, when you were a child, were there more conversations in the streets when you were a child?'
Participant: 'Of course, *there were*$_{Sing}$ more.'

2.1.3 Hearer-old entities newly instantiating a variable

The third type Ward & Birner (1995) identify is that of hearer-old entities newly instantiating a variable, resulting in a list reading. According to these authors, list-reading definites require a context that evokes an open proposition of the type 'X is an element of the category Y' (Ward & Birner, 1995: 734-735). The use of a presentational expression, then, is motivated by the fact that the elements of the list are presented as hearer-new instances of the category. In turn, the use of the definite article is licensed by the fact that hearers are expected to uniquely identify the list items (Suñer, 1982: 88-90; Ward & Birner, 1995: 734-735). This is shown in example (32), in which the character Teófilo Huamani first evokes the category *gran civilización* 'great civilization', upon which the other character enumerates some examples with a pluralized presentational *haber* expression. Similar cases can be found with the singular construction, as is shown in example (33).

(32) Teófilo Huamani: Porque, a mí, los balcones representan la opresión.
Professor Brunelli: ¿Se puede saber a quién o a qué oprimen estos pobres balcones?
Teófilo Huamani: Antes de que llegaran aquí los forasteros que los trajeron, en el Perú *había* una gran civilización, profesor.
Professor Brunelli: La de los incas, lo sé muy bien. Y, antes, *habían* los chimús, los nazcas, los tiahuanacos, muchos más (Real Academia Española, 2008b-, Theater, Peru).
Teófilo Huamani: 'Because, for me, the balconies represent oppression.'
Professor Brunelli: 'Can I know who or what these poor balconies oppress?'
Teófilo Huamani: 'Before the foreigners came here that brought them, in Peru *there was* a great civilization, professor.'
Professor Brunelli: 'That of the Incas, I know it very well. And before, *there were*$_{Plur}$ the Chimus, the Nazcas, the Tiahuanacos, many more.'

(33) Bueno, yo creo, francamente, que tenemos que adaptarnos, en nuestro teatro, y en otras manifestaciones de la vida puertorriqueña, a la forma de hablar puertorriqueña.
…
Hay también el yeísmo. ¿Verdad? Pero el yeísmo está aceptado ya en el resto de Hispanoamérica, y *hay* los apócopes de la ese final, que ocurren mucho en, en Puerto Rico (Real Academia Española, 2008b-, Spoken, Puerto Rico).
'Well, frankly, I think that we have to adapt ourselves, in our theater, and in other manifestations of Puerto Rican life, to the Puerto Rican way of talking.
…
There is also the yeismo. Right? But the yeismo is already accepted in the rest of Hispanic America and *there are*$_{Sing}$ the apocopes of word-final *s*, which occur a lot in, in Puerto Rico.'

In this light, one would expect proper names without determiners also to occur in lists. Indeed, in English, this is possible, as is shown by the felicity of the co-occurrence of the proper name *John McCain* with presentational *there is/there are* in example (34).

(34) I think, one, *there is* John McCain and there is everybody else (Davies, 2008-, Press).

However, this is impossible with presentational *haber*. Suñer (1982: 82) attributes this to the fact that a proper name would require differential object marking with the preposition *a* (e.g., *a John McCain*), which *haber* does not allow, because it lacks an (agentive) subject (Delbecque, 2002: 107; Torrego-Salcedo, 1999: 1785, 1794-1795).

Finally, it should also be observed that definites can be licensed in multiple ways. Consider, for instance, example (35). Here, the definites *los juegos estos que te digo* 'the games that I told you about' and *la televisión* 'the television' are licensed as uniquely identifiable elements of the category *pasatiempos* 'pastimes' evoked in the interviewer's question. At the same time, however, the use of the first definite NP is also licensed as a reminder.

(35) Interviewer: ¿Este, y cuando tú eras niña, qué pasatiempos habían?
Participant: ¿Qué pasatiempos entonces habían? E, t, bueno, *habían* los juegos estos que te digo, la televisión, aunque eran pocos los muñequitos que habían, pero, m, m, pero habían algunos (LH03M12 /LH287).
Interviewer: 'Er, when you were a girl, what pastimes were there?'
Participant: 'What pastimes were there? Er, t, well, *there were*$_{Plur}$ those games that I told you about, the television, although the cartoons that there were, were few, but, m, m, but there were some.'

2.1.4 Hearer-new entities with uniquely identifying descriptions

The nominal argument can also be marked by a definite determiner because it introduces a hearer-new entity with a uniquely identifying description (Lakoff, 1987: 546; Ward & Birner, 1995: 735-736). Contrary to what we have seen for the other types of definites, the degree of acceptability of this type does not hinge upon the context. Rather, this sort of definite argument introduces brand-new information, but in such a way that the hearer can immediately identify the unique referent the speaker is talking about, which licenses the use of definite determiners (Abbott, 2004: 136; Langacker, 1991: 98). Example (36) shows that this interpretation emerges with NPs introduced by demonstratives and followed by adnominal descriptions. Cases like this can also be found with singular presentational *haber* (see example 37), as well as with the definite article (see examples 38 and 39).

(36) No *habían* esos bares y esos, s, esas cosas que *hay*, que están creando problemas (SD15M21/RD1810).
'*There weren't*$_{Plur}$ these bars and these, t, these things that *there are*$_{Sing}$, that are creating problems.'

(37) Yo creo que el país tiene otras urgencias ahora y que, en la situación en que está, la cultura no es una de ellas. Cada quien irá haciendo lo que pueda con los pocos recursos que tenga.
…
Ya no *hay* esos grandes subsidios o esas grandes exposiciones y puestas en escena que podías traer del exterior (Real Academia Española, 2008b-, Press, Venezuela).
'I think that the country has other urgencies now and that, in the situation in which it is, the culture is not one of them. Everyone will keep doing what they can with the few resources they have.
…
There aren't$_{Sing}$ these large subsidies or these large exhibitions or stagings that you could bring from abroad anymore.'

(38) Ante las denuncias de robo, acoso sexual y amenazas que *han habido* en el recinto de Río Piedras de la Universidad de Puerto Rico (UPRRP), el rector Carlos Severino, aseguró que "está cambiando la manera en que hacen la seguridad", al incluir mayor patrullaje y retomar el tema de la acreditación de la Guardia Universitaria (Internet, Press, Puerto Rico, http://goo.gl/oiBzR3).
'Faced with the allegations of theft, sexual harassment, and threats that *there have been*$_{Plur}$ at the Rio Piedras campus of the University of Puerto Rico (UPRRP), the Rector Carlos Severino, assured that "He is changing the way they do security" by including more patrols and revisiting the issue of the accreditation of the University Guard.'

(39) Los insectos han sobrevivido cuatro de las cinco grandes extinciones que *ha habido* en el planeta (Internet, Press, Puerto Rico, http://goo.gl/K4ze4g).
'Insects have survived four of the five great extinctions that *there have been*$_{Sing}$ on the planet.'

Uniquely identifying descriptions may also be constructed with an anaphoric pronoun followed by a restrictive relative clause (RAE & ASALE, 2009: §15.6r), as in example (40). Since this type of expression involves accusative pronouns, they are less common with pluralized presentational *haber*. Still, some cases can be documented, as is shown in example (41).

(40) Bueno, nosotros en Cuba les llamamos 'guaguas'. Son unos pequeños insectos, no recuerdo de qué familia. Los *hay* blancos, pequeñitos, los *hay* que parecen cucarachitas pequeñitas que específicamente succionan, chupan los jugos vegetales y, entonces, esto lógicamente empobrece las plantas y hay que luchar contra ellas (Davies, 2002-, Spoken, Havana).
[Well, we in Cuba, we call them 'guaguas'. They are little insects, I don't recall of which family. Them$_{Acc}$ *there are*$_{Sing}$ white, really small, them$_{Acc}$ *there are*$_{Sing}$ that look like little cockroaches, which specifically suction, suck the plant's juices and, well, this logically weakens the plants and you have to fight against them.]
'Well, we in Cuba, we call them 'guaguas'. They are little insects, I don't recall of which family. *There are*$_{Sing}$ some small white ones, *there are*$_{Sing}$ some that look like little cockroaches, which specifically suction, suck the plant's juices and, well, this logically weakens the plants and you have to fight against them.'

(41) Claro, al haber tantos alumnos en el aula, un solo maestro para cuarenta o cincuenta muchachos. Y, entonces, los *habían* disciplinados, pero los *habían* que eran la candelita (LH21H11/LH2835-LH2836).
[Of course, with there being so many pupils in the classroom, a single teacher for forty or fifty kids. And, so, them$_{Acc}$ *there were*$_{Plur}$ disciplined, but them$_{Acc}$ *there were*$_{Plur}$ who were a handful.]
'Of course, with there being so many pupils in the classroom, a single teacher for forty or fifty kids. And, so, *there were*$_{Plur}$ disciplined ones, but *there were*$_{Plur}$ some who were a handful.'

Further examples of this type can be found with superlatives (see examples 42 and 43) and cataphoric-reference NPs (see examples 44 and 45) (Bolinger, 1977: 117-118; Suñer, 1982: 80, 82-84; Ward & Birner, 1995: 737).

(42) Por plata *han habido* los más extraños cambios de postura en toda la historia humana (Internet, Message board, Chile, http://goo.gl/nuyRpY).
'For money *there have been*$_{Plur}$ the strangest posture changes in the entire human history.'

(43) Creo que en el Pri *hay* los mejores políticos del Estado, las gentes que tienen la mejor experiencia de gobierno (Davies, 2002-, Spoken, Mexico).
'I think that in the Pri Party *there are*$_{Sing}$ the best politicians of the state, the people that have the best governance experience.'

(44) Es posible que la ola que decayó, a nuestro juicio, en esos días, se hubiese elevado de nuevo, si nosotros convocamos a la huelga y la anunciamos 48 horas antes. Claro, *habían* los criterios siguientes. Si nosotros anunciamos la huelga, el ejército, el régimen, toma medidas en una serie de puntos que nos interesa atacar (Internet, Magazine, Cuba, http://goo.gl/hY5zEh).
'It is possible that the wave that fell, in our view, in those days, would have raised itself again, if we called for a strike and announced it 48 hours in advance. Of course, *there were*$_{Plur}$ the following criteria. If we announce the strike, the army, the regime takes action on a number of points we are interested in attacking.'

(45) En los congresos de París (1989) y de Roma (1993) se presentaron numerosos estudios multidisciplinares y todos confirmaban la antigüedad del lienzo. Entre ellos *hay* los dos siguientes: 1. La irradiacion.
…
2. Los incendios (Internet, Website, Peru, http://goo.gl/VLIh3P).
'At the conferences of Paris (1989) and Rome (1993), numerous multidisciplinary studies were presented and all confirmed the antiquity of the cloth. Among them *there are*$_{Sing}$ the following two: 1. Radiation
…
2. Fires.'

The last type of uniquely identifiable definite nouns is known as 'containing inferables' (Prince, 1992: 303-305). With this type, the reference of the noun is inferred from its adnominal modifier. For instance, in the pluralized case presented in example (46), the reference of *las partes* 'the parts' can be inferred from the PP *de un hombre* 'of a man'. Similarly, in the singular example provided in (47), the specific meaning of *los elementos* 'the elements' is inferable from *de un golpe de Estado* 'of a (typical) coup'.

(46) En el asiento trasero estaban dos bolsas de plástico color negra, en cuyo interior *habían* las partes de un hombre de aproximadamente 40 años de edad (Internet, Press, Mexico, http://goo.gl/m0zfV8).
'Two black-colored plastic bags were in the back seat, inside which *there were*$_{Plur}$ the parts of a man of about 40 years of age.'

(47) Les expliqué que *había* los elementos de un golpe de Estado (Internet, Press, Argentina, http://goo.gl/lW0Rfo).
'I explained to them that *there were*$_{Sing}$ the elements of a (typical) coup.'

With English presentational *there is/there are,* Ward & Birner (1995: 737) find that containing inferables can only be used felicitously when a 'conventional relationship'

holds between the entities denoted by the head (*las partes* 'the parts' and *los elementos* 'the elements', in the examples) and the modifier (*de un hombre* 'of a man' and *de un golpe de Estado* 'of a coup' in the examples).[10] Indeed, for many containing inferables, an 'intrinsic metonymic association', such as the part-whole relationships in the examples (Croft & Cruse, 2004: 216-217), can be identified between the modifying adnominal and the noun.

In contrast, when no metonymic association exists between the noun and its modifier, the use of a definite determiner leads to an infelicitous expression, as is evident from example (48) (Ward & Birner, 1995: 738).

(48) **There was* the picture of a young black couple among his papers (constructed example. From Ward & Birner, 1995: 738). [11]

Additionally, because the information status of a containing inferable definite NP is determined by that of the adnominal modifier (Birner, 1994: 252), we can expect its use in a presentational *haber* expression to be odd when the PP is hearer-old. Judging from the modified versions of examples (46) and (47) cited in (49) and (50), this appears to be the case. [12]

(49) *En el asiento trasero estaban dos bolsas de plástico color negra, en cuyo interior *habían* las partes del hombre de aproximadamente 40 años de edad (constructed example).
*'Two black-colored plastic bags were in the back seat, inside which *there were*$_{Plur}$ the parts of the man of about 40 years of age.'

[10] See RAE & ASALE (2009: §15.6s) for a similar analysis.

[11] However, if this expression were to occur in a context that construes it as another type of definite NP, it would be fine. For example, *Then, there was the picture of a young black couple among his papers* evokes a list reading.

[12] When we add a restrictive relative clause to the PP, these utterances are well-formed, as is evident from examples (i) and (ii).

(i) En el asiento trasero estaban dos bolsas de plástico color negra, en cuyo interior *habían* las partes del hombre de aproximadamente 40 años de edad que habían estado buscando (constructed example).
'Two black-colored plastic bags were in the back seat, inside which *there were*$_{Plur}$ the parts of the man they had been looking for.'

(ii) Les expliqué que *había* los elementos del golpe de Estado que ya denunció mi antecesor en el cargo (constructed example).
'I explained to them that *there were*$_{Sing}$ the elements of the coup my predecessor had already denounced.'

However, examples (i) and (ii) are not containing inferables, because the hearer does not draw on the metonymical association between the PP and the noun to identify the referent of the noun, but rather uses the information provided by the PP and its restrictive relative clause.

(50) *Les expliqué que *había* los elementos del golpe de Estado (constructed example).
*'I explained to them that *there were*$_{Sing}$ the elements of the coup.'

2.1.5 False definites

Finally, definite NPs can be used to refer to brand-new referents,[13] in which case the definite determiners function as indications of intensity (Suñer, 1982: 81; Ward & Birner, 1995: 739). Since the information contained within these NPs is truly new, false definites can be used freely with both English presentational *there is/there are* (Ward & Birner, 1995: 738-740) and Spanish presentational *haber*. With singular NPs, this interpretation may emerge in Spanish with certain superlatives (see example 51), the close-to-hearer deictic *esa/ese* 'this' (see example 52), and the definite article (see example 53). With plural NPs, false definite readings may arise with *todas las*/*todos los* 'all the' (see examples 54 and 55) and the less common form *cuantas/cuantos* 'all the' (see examples 56 and 57) (RAE & ASALE, 2009: §15.6k, §19.3a; Suñer, 1982: 81).

(51) Aun a riesgo de repetirme les quiero decir que no *hay* el menor problema y que los ciudadanos de Canarias pueden estar tranquilos, igual que los ciudadanos de toda España (Real Academia Española, 2008b-, Press, Canary Islands).
'Even at the risk of repeating myself I want to say to you that *there isn't* the slightest problem and that the Canarian citizens can rest assured, just like the citizens across Spain.'

(52) En mi casa yo también, e, yo soy, éste era un pobre, no, muy pobre, y, entonces, no había esa, no *había* ese dinero para tener unos juguetes nuevos, así constantemente, tener muchos juguetes (LH01H22/LH41).[14]
'At home I as well, er, I am, this was a poor fellow, right, really poor, and, well, there wasn't this, *there wasn't* this money to have some new toys, like that, constantly, to have a lot of toys.'

(53) *Hay* el hombre y *hay* la mujer. Y cada uno tiene cosas distintas (René Marqués, *La Mirada*. From RAE & ASALE, 2009: §15.6p).
'*There is* man and *there is* woman. And both have different things.'

[13] Referents that have not been evoked in earlier discourse and with which the hearer has yet to establish mental contact (Prince, 1992: 318).

[14] I collected the examples of presentational *haber* + singular NP specifically for this chapter. As stated in Chapter 4.3, only the cases of *haber* + plural NP will be considered in the quantitative analyses of Chapter 6.

(54) Cuentan los abuelos que los tres campesinos se perdieron en aquella espesa y misteriosa selva, llegando, según la leyenda, al encanto invisible de doña Ñuisa, un paraje mítico encantado perdido en la selva, donde *habían* todos los frutos que hay sobre la tierra, hermosos jardines, quebradas de aguas cristalinas, con arenas de plata y piedras de oro (Internet, Blog, Colombia, http://goo.gl/sjMLgx).
'The grandfathers recount that the three farmers got lost in that thick and mysterious forest, reaching, as the legend goes, the invisible charm of Mrs. Ñuisa, a mythical enchanted place, lost in the woods, where *there were*$_{Plur}$ all the fruit trees that there are on earth, beautiful gardens, waterfalls with crystal-clear water, with silver sands and golden stones.'

(55) Halló aquí belleza y pobreza, pero también un pueblo alerta, descalzo y sensible, de infatigables manos hacedoras, dueño de una tierra inmensa donde *hay* todos los paisajes y los climas, los frutos y los sueños (Real Academia Española, 2008b-, Fiction, Mexico).
'He found here beauty and poverty, but also an alert, barefooted, and sensible people, with untiring working hands, owner of an immense land, where *there are*$_{Sing}$ all the landscapes, the climates, the fruits, and the dreams.'

(56) Habían mangos, habían piñas, *habían* cuantas frutas había (SJ16H21/SJ1951).
'There were mangoes, there were pineapples, *there were*$_{Plur}$ all the fruits *there were*$_{Sing}$.'

(57) En fin abrí el cajón del buró para buscar el control de la televisión, ahí *había* cuantas cosas extraordinariamente desordenadas pudiese imaginar: pulseras, coletas, trabas, collares entre cientos y cientos de cosas (Internet, Blog, Chile, http://goo.gl/UYLGw9).
'Eventually, I opened the drawer of the desk to look for the television remote control, in there, *there were*$_{Sing}$ all the extremely messy things I could imagine: bracelets, pigtails, hair clips, necklaces, among hundreds and hundreds of things.'

In sum, in this section I have shown that the nominal arguments of pluralized and singular presentational *haber* are interpreted as being present in a stative situation. This suggests that the NP of both constructions is assigned a zero argument role. Additionally, the examples presented in this section suggest that Ward & Birner's (1995) analysis of English presentational expressions is also valid for pluralized and singular presentational *haber*. This points to a shared pragmatic constraint on the nominal arguments of both presentational *haber* constructions, namely, that it has to convey new information to the hearer. As we will see in the next section, this constraint explains why the nominal argument of pluralized presentational *haber*

typically fails syntactic tests of subjecthood (e.g., Rodríguez-Mondoñedo, 2006; Suñer, 1982).

2.2 Syntactic properties

In discussions of the syntactic status of the nominal argument of presentational *haber* (e.g., Gómez-Torrego, 1994: 30; Rodríguez-Mondoñedo, 2006: 334; Suñer, 1982: 22), it has often been observed that subject-marked personal pronouns are barred from appearing with both the pluralized and the singular construction, as is shown in examples (58) and (59). Contrary to what is claimed by the main hypothesis, this would suggest that the NP invariantly behaves as an object (Rodríguez-Mondoñedo, 2006: 334). In order to shed more light on this matter, in this section, I will review the argumentation that has been proposed in favor of this position. This will lead to the conclusion that most object-like characteristics of the nominal of presentational *haber* can actually be traced back to the information-status constraint identified in the previous section.

(58) **Había* ellos (constructed example).
*'*There were*$_{Sing}$ they$_{Nom}$.'

(59) **Habían* ellos (constructed example).
*'*There were*$_{Plur}$ they$_{Nom}$.'

To start with, drawing on Keenan's (1976) list of subject properties, Rodríguez-Mondoñedo (2006: 330) and Suñer (1982: 121) indicate that the NP of pluralized and singular presentational *haber* cannot remain implicit in coordinated structures, whereas this is usually possible for subjects, as is evident from Suñer's (1982) examples cited in (60).

(60) a. *Irradiaban* luz y *olían* agradablemente dos docenas de rosas (constructed example. From Suñer, 1982: 104).
'*Radiated* light and *smelled* pleasantly two dozens of roses.'
b. **Habían* y *olían* agradablemente dos docenas de rosas (constructed example. Adapted from Suñer, 1982: 104).
*'*There were*$_{Plur}$ and *smelled* pleasantly two dozens of roses.'
c. **Había* y *olían* agradablemente dos docenas de rosas (constructed example. From Suñer, 1982: 104).
*'*There were*$_{Sing}$ and *smelled* pleasantly two dozens of roses.'

Second, it is impossible to interpret the NP of *haber* as coreferential with the subject of a matrix verb (Gómez-Torrego, 1994: 30; Rodríguez-Mondoñedo, 2006: 330-331; Suñer, 1982: 20), as is shown in example (61).

(61) a. *Los perros *quieren haber* en el jardín (constructed example).
*'The dogs *want to there-be*$_{Plur}$ in the garden.'
b. *Los perros *quiere haber* en el jardín (constructed example).
*'The dogs *want to there-be*$_{Sing}$ in the garden.'

Third, the default word order with presentational *haber* is verb + NP. Since Spanish, and especially Caribbean Spanish, is mainly an SVO-language (Aponte-Alequín, 2014: 182-183; Morales, 1982, 1989, 1997, 1999; Silva-Corvalán, 2001: 171; Suñer, 1982: 281), this ordering can be interpreted as evidence in favor of the object status of the NP (Montes de Oca, 1994: 11). Moreover, placing the nominal in subject position[15] leads to an unacceptable expression, as is evident from examples (62) and (63).

(62) *Unos hombres *habían* en el jardín (constructed example. Adapted from Rodríguez-Mondoñedo, 2006: 333).
*'Some men *there were*$_{Plur}$ in the garden.'[16]

(63) *Unos hombres *había* en el jardín (constructed example. Adapted from Rodríguez-Mondoñedo, 2006: 333).
*'Some men *there were*$_{Sing}$ in the garden.'

However, Givón (1999: 94-96) has shown that the subject properties proposed by Keenan (1976) are to a large extent epiphenomena of the tendency for subjects to have discourse-old information status. Indeed, the infelicity of nominative personal pronouns can be explained by the fact that these require hearer-old/discourse-old information status (Bolinger, 1977: 91). Also, as we will see below, the impossibility to use an implicit nominal argument in coordinated structures shows that the use of implicit arguments requires that the implicit portion of the event frame has already appeared in discourse (Goldberg, 2006a: 190). The word order that is typically displayed by presentational *haber* expressions is also predictable from the information status of the NP argument, as cross-linguistically, new information tends to be placed in post-verbal position (e.g., Birner, 1994; Birner & Ward, 1996). In turn, the fact that the nominal argument of the presentational *haber* constructions cannot be interpreted as coreferential with the subject of a matrix verb illustrates that the NP can only be interpreted as a zero participant.

In other words, when the pragmatics and the semantics of presentational *haber* constructions are taken into account, syntactic tests do not necessarily prove that the nominal argument also functions as an object with pluralized presentational *haber*.

[15] That is, at the beginning of the utterance.

[16] Rodríguez-Mondoñedo (2006: 333) translates his example *Un hombre había en el jardín* as 'A man was in the garden'. However, this is not a presentational, but rather a locative expression, which answers to different information-status constraints and would be translated in Spanish with the verb *estar*. In English, the conceptual import of presentational *haber* can only be rendered correctly by *there is*/*there are* constructions.

Rather, they show that only the absence/presence of verb agreement can be taken as a formal clue for the grammatical status of the NP. Let us turn now to the adverbial phrase.

3. The adverbial phrase

As was already mentioned in section 1, Lakoff (1987: 490) describes the meaning of English presentational *there is/there are* as an ICM that introduces a new referent into discourse while situating it in a mental space. The adverbial phrase of the English presentational *there is/there are* construction (e.g., *In the U.S.* in example 64) is the element that sets up this mental space (Lakoff, 1987: 542-543).

(64) In the U.S., *there are* now more jobs in the wind industry than in the entire coal industry (Davies, 2008-, Magazine)

Similarly, with Spanish presentational *haber,* the adverbial phrase creates the mental space in which the constructions locate the referents of their nominal arguments (see Hernández-Díaz, 2006: 1130-1132; Lyons, 1967; Meulleman & Roegiest, 2012). Syntactically, this implies that the presence of the adverbial expression cannot be considered optional (Meulleman & Roegiest, 2012: 68-69). Rather, the fact that the adverbial contains necessary information for the interpretation of the expression suggests that its syntactic status is that of an 'obligatory adjunct'(Goldberg & Ackerman, 2001), that is, a profiled adverbial phrase.

Because the adverbial phrase is not claimed to refer to a physical location, but rather serves to construct "small conceptual packets … for the purpose of local understanding and action" (Fauconnier & Turner, 1996: 113), it is no surprise that they may be of a spatial (see examples 65 and 66), a temporal (see examples 67 and 68) (e.g., Clark, 1978: 89; Hernández-Díaz, 2006: 1130-1132; Lyons, 1967; Meulleman & Roegiest, 2012) or another nature.

(65) En el Norte de Italia *habían* muchas guerrillas (SD16H22/RD2188).
'In the North of Italy, *there were*$_{Plur}$ a lot of guerrillas.'

(66) Se abría en octubre la universidad, pero yo venía en enero, porque en casa no *había* cuartos para asistir desde, el año entero (SD16H22/RD2068).
'The university opened in October, but I came in January, because at home *there weren't*$_{Sing}$ moneys to attend from, the entire year.'

(67) E, en los, en los tiempos de antes, la, no *habían* tantas leyes (SJ16H21/SJ1971).
'Er, in the, in the old days, the, *there weren't*$_{Plur}$ so many laws.'

(68) Es, e, en estos momentos tiene sala, comedo, sala-comedor, la cocina, el baño y, bueno, tres cuartos que antes no existían. Antes solamente *había* dos (LH07M11/LH842).
'It is, er, at the moment it has a living room, a dinin, a living-dining room, the kitchen, the bathroom, and, well, three bedrooms that didn't exist before. Before, *there were*$_{Sing}$ only two.'

Additionally, this characterization of the adverbial phrase of the presentational *haber* constructions correctly predicts that they need not be made explicit when the hearer and the speaker can be expected to be able to recover the mental space from context (Goldberg, 1995: 58-59). This is the case when the expression situates the NP in a previously constructed mental space[17] (see examples 69 and 70) or the current base space, as in examples (71) and (72).

(69) Interviewer: ¿Este, y cuando tú eras niña quién te cocinaba?
Participant: Mi abuela.
Interviewer: ¿Tu abuela? ¿Era co, era buena cocinera?
Participant: Mediana. O sea, pues, o sea, *habían* cosas que las hacía muy bien, pero otras cosas que no (SD24M12/RD3202).
Interviewer: 'Er, and when you were a child, who cooked for you?'
Participant: 'My grandmother.'
Interviewer: 'Your grandmother? Was she a co, was she a good cook?'
Participant: 'Average. That is, well, that is, *there were*$_{Plur}$ things that she made very well, but other things that she didn't.'

[17] Mental spaces can be set up by adverbial elements, phrases and clauses, but also by verb tenses, negation, or matrix verb constructions such as, for example, *to believe that, to remember that, to think that* (Croft & Cruse, 2004: 32-39).

(70) La primera experiencia que yo recuerdo fue el huracán Hugo, e, que azotó a la Isla, e, prácticamente el Área Metropolitana fue la más im, e, impactada. El Área Sur como te dije, yo viví en Ponce, yo estaría como en el noveno grado, octavo-noveno no recuerdo el año. E, fue en los ochenta, e, pero recuerdo que fue un, un huracán bastante fuerte. ¿Qué categoría? No me preguntes, pero fue bastante fuerte. Recuerdo que, e, *hubo* mucha lluvia y muchas inundaciones (SJ06H12/SJ767).
'The first experience that I remember was the hurricane Hugo, er, which struck the Island, er, practically, the Metropolitan Area was the most, er, impacted. The Southern Area, like I told you, I lived in Ponce, I would have been in the ninth grade, eighth-ninth I don't remember the year. Er, it was in the eighties, er, but I remember that it was a, a pretty tough hurricane. What category? Don't ask, but it was pretty rough. I remember that, er, *there were*$_{Sing}$ a lot of rain and many floods.'

(71) Pero, este, sí, *hayn* platos como que es, específicos de diciembre (SJ05M12/SJ655).
'But, er, yes, *there are*$_{Plur}$ dishes, like, spe, typical of December.'

(72) *Hay* veces que sí lo notas, porque ellos te usa, te usan unas palabras que te arrastran la ere, '*jj*' la hacen como, y tú te das cuenta ahí que no es de San Juan (SJ01M22/SJ68).
'*There are*$_{Sing}$ times that you do notice it, because they use, they use some words where they drag along the *r*, '*jj*' they do it like, and there you immediately realize that they're not from San Juan.'

When it comes to the information status of the adverbial phrase, in my corpus, all possible combinations between Prince's (1992) hearer- and discourse-oriented levels of information status seem to occur. For instance, in example (73) the speaker first introduces a particular bridge into discourse and uses it later on in the interview to situate *muchos muchachos bonitos* 'many pretty boys', for which the adverbial has discourse-old/hearer-old information status.

(73) Pues, yo tengo miedo a las alturas. T, y estuve en un campismo, que todo el mundo, t, se tiraba de <u>un puente</u>. Todo el mundo, el mundo se lanzaba de <u>un puente</u>, como de, o sea, seis metros así de ver. Y yo era adolescente en esa época. Y <u>en el puente</u> *habían* muchos muchachos bonitos (LH09M12/LH1168).
'Well, I have a fear of heights. T, and I was on a camping, where everyone, t, was jumping off <u>a bridge</u>. Everyone, everyone was jumping of <u>a bridge</u>, of like, that is, six meters, judging from sight. And I was a teenager at the time. And <u>on the bridge</u> *there were*$_{Plur}$ many pretty boys.'

Discourse-new/hearer-old adverbial phrases can also be found. These usually refer to geographic landmarks or areas the hearer is expected to know, such as, for instance, the city of San Juan de la Maguana in example (74), or the Pinar del Río province in example (75).

(74) E, en San Juan no *habían* escuelas privadas (SD04M22/RD447).
'Er, in San Juan, *there weren't*$_{Plur}$ private schools.'

(75) Sí, *hay* lugares bellos en Pinar del Río (LH21H11/LH2868).
'Yes, *there are*$_{Sing}$ beautiful places in Pinar del Río.'

Finally, the adverbial phrase can refer to a mental space that is both new to the hearer and to discourse. For instance, in examples (76) and (77), *en una escuela secundaria* 'in a secondary school' and *en lugares públicos* 'in public places' are newly introduced into the conversation. Still, the hearer can be expected to build up the specific indefinite mental space the speaker has in mind.

(76) Allá en Estados Unidos cuando yo, yo trabajé en una escuela prima, en una escuela secundaria en Nueva York y *habían* muchachos de distintas partes de Latinoamérica (SD24M12/RD3212).
'Over there in the United States, when I, I worked in a pri, in a secondary school in New York and *there were*$_{Plur}$ kids of different parts of Latin America.'

(77) Interviewer: ¿Este, y, entonces, los amigos se, se reunían y se, se co, se hablaban o…?
Participant: M, se hablaban o preparaban un motivo, un, un, una especie de fiesta, porque siempre era mejor reunirse en, en una casa que, no en lugares públicos, que *hay* otros riesgos (LH16H22/LH2206).
Interviewer: 'Er, and, well, the friends got to, together and they, ta, they talked to each other or…?'
Participant: 'M, they talked to each other or they prepared an occasion, a, a, a sort of party, because it was always better getting together in a home than, not in public places, where *there are*$_{Sing}$ other risks.'

In any case, the variety of configurations documented in the corpus suggests that the presentational *haber* constructions do not specify the information status of this slot. Let us now consider the conditions that constrain the use of implicit adverbial phrases and/or nominal arguments.

4. Implicit nominal arguments and adverbial phrases

As noted in section 2.1 the POINTING-OUT ICM implies that without proper context, the NP argument cannot remain implicit, as it carries virtually the entire conceptual import of the clause (Goldberg, 2005a: 29, 2005b: 232). Nevertheless, under specific discourse conditions, it need not be made explicit again. As observed in section 2.1.2, in my corpus this is especially common for indefinite NPs that introduce hearer-new tokens of hearer-old types, as shown in example (78).

(78) Participant: Niños en la calle, yo creo.
Interviewer: ¿Antes no?
Participant: Hay más.
Interviewer: ¿No habían tantos?
Participant: *Habían,* siempre *han habido* (SD19M12/RD2513-RD2514).
Participant: 'Children on the streets, I think.'
Interviewer: 'Before not?'
Participant: 'There are more.'
Interviewer: 'There weren't as many?'
Participant: '*There were*$_{Plur}$*, there have* always *been*$_{Plur}$.'

Similarly, the fact that the adverbial phrase is profiled and sets up a mental space in which presentational *haber* locates the referent of the NP implies that the adverbial can only be omitted felicitously when it is recoverable (Goldberg, 1995: 58-59, 2006a: 39), that is, when it refers to the base space or a previously evoked mental space. Isolated examples such as (79) and (80), which leave us wondering against which setting we have to interpret the utterances, suggest that this is the case.

(79) *Podrían haber* días en que yo tenía dos horas libres entremedio (SJ13H11/SJ1566).
'*There could be*$_{Plur}$ days that I had two hours of free time in between.'

(80) Claro, sí *hubo* muertos (SD20H12/RD2682).
'Of course, yes, *there were*$_{Sing}$ casualties.'

In contrast, examples such as (81) and (82) are conceptually complete, because they locate the referent of the NP in the current base space. Let us now resume the most important results of this chapter.

(81) *Habrán* gentes que lo hagan (SD05H11/RD594).
'*There will be*$_{Plur}$ people that do it.'

(82) El racismo muchas veces viene porque muchos blancos ignoran de que *hay* blancos que s, negros que son tesoros (SD05H11/RD562).
'Often, racisms comes because many whites are unaware that *there are*$_{Sing}$ whites that a, blacks that are treasures.'

5. Summary and box diagrams

In this chapter, I have argued that both the pluralized and the singular variant of the presentational construction with *haber* encode the POINTING-OUT ICM proposed by Lakoff (1987). The two constructions also assign the same zero argument role to their nominal argument, which functions as the trajector of the clause. Since alternations like this serve to provide speakers with different ways to package information (Goldberg, 2005a: 37, 2005b: 236), subsequently, I have investigated whether pluralized and singular presentational *haber* display any differences in this respect. This has shown that the nominal argument of both variants of the presentational *haber* construction has to provide new information to the hearer. Then, I have indicated that due to this information-status constraint, the results of syntactic tests do not necessarily prove that the nominal is always a direct object. Rather, the only formal clue that remains is the absence or presence of verb agreement. Therefore, this chapter has shown that it is at least possible that pluralized presentational *haber* has a subject, as the main hypothesis claims.[18]

Additionally, I have demonstrated that the two variants of the presentational construction with *haber* include a profiled adverbial phrase, which functions as setting and evokes the mental space in which the construction localizes the referent of the nominal argument. Subsequently, it was shown that under certain discourse conditions, both the nominal argument and the adverbial phrase may remain implicit. For the nominal argument, this results in its interpretation as a hearer-new token of a hearer-old type, a reading that also emerges for definites in certain discourse contexts. For the adverbial phrase, this results in its interpretation as referring either to the base space or to a previously established mental space.

In sum, the data presented in this chapter suggest that it is not *a priori* impossible that the pluralization of presentational *haber* constitutes a competition between two construction schemas (see Figure 5.1 and Figure 5.2) that only differ when it comes to the syntactic function of their nominal arguments and the social types associated to their relative frequencies. In Chapter 3.1.7, the claim was made that this type of meaning is represented mentally as a probability linking distributions of alternating forms with social types. In order to capture this, besides the pragmatic, semantic, and syntactic constraints that were discussed in this chapter, Figure 5.1 and Figure 5.2 also

[18] See Chapter 3.3.1.

include a social layer. In the next chapter, after exploring the cognitive factors that shape the variation between the presentational *haber* constructions and their implications for the main hypothesis, the social distribution of the alternation will be investigated.

Figure 5.1: The singular presentational *haber* construction

Sem	POINTING-OUT	<	**location**	**zero**	>
R: instance	|R		|	|	
	haber	<	**setting**	**participant**	>
	↓		↓	↓	
Syn	V		adverbial phrase	object	
Prag			-	hearer-new	
Soc					

Figure 5.2: The pluralized presentational *haber* construction

Sem	POINTING-OUT	<	**location**	**zero**	>
R: instance	|R		|	|	
	haber	<	**setting**	**participant**	>
	↓		↓	↓	
Syn	V		adverbial phrase	subject	
Prag			-	hearer-new	
Soc					

6

A Cognitive Construction Grammar approach to the pluralization of presentational haber *and its social distribution*

The preceding chapters have laid the basis for a comprehensive analysis of *haber* pluralization in Caribbean Spanish. Against their background, this chapter will evaluate whether the pluralization of presentational *haber* can be described as an ongoing language change from below involving a competition between the two variants of the presentational *haber* construction presented in Chapter 5.5. Particularly, after reviewing the general distribution of pluralized and singular presentational *haber* (see section 1), section 2 will focus on the effect of the three cognitive factors that were described in Chapter 3.3.2 (markedness of coding, statistical preemption, and structural priming). Subsequently, in section 3, the social distribution of *haber* pluralization will be investigated. The chapter concludes with a summary of the results (see section 4).

1. General distribution

Table 6.1 shows that, using the methods described in Chapter 4, a total of 5,589 tokens of presentational *haber* followed by a plural NP have been collected. Of these, across the communities, 44.3% (N=2477/5589) correspond to the pluralized presentational *haber* construction. For the individual speech communities, the total number of tokens ranges between about 1,650 to more than 2,000. The rates of *haber* pluralization range between 41.3% and 46.7%.

Table 6.1: Pluralized and singular presentational *haber* in the Spanish of Havana, Santo Domingo, and San Juan

	Havana		*Santo Domingo*		*San Juan*	
	N	*%*	*N*	*%*	*N*	*%*
Pluralized	934	44.6	859	46.7	684	41.3
Singular	1159	55.4	982	53.3	971	58.7
Total	2093	100	1841	100	1655	100

In other words, Havana, Santo Domingo, and San Juan are not markedly different from one another when it comes to their overall rates of *haber* pluralization. In contrast, when compared to earlier work, the rates of *haber* pluralization reported here are far lower. That is to say, the studies reviewed in Chapter 2.2 typically document pluralized *haber* in about 60% of the cases (Bentivoglio & Sedano, 1989; D'Aquino-Ruiz, 2004, 2008; Díaz-Campos, 2003). In San Cristóbal de Los Andes, San Salvador, and Valencia de Venezuela, pluralized *haber* even occurs in around 80% of the cases (Freites-Barros, 2008; Navarro-Correa, 1992; Quintanilla-Aguilar, 2009). Yet, the differences between the figures reported here and the findings of earlier investigations appear to be due to the fact that this study includes the variation between the present-tense form *hay* and its rather infrequent vernacular plural *hayn*.[1] Without these two forms, the frequency of pluralized *haber* rises to 60.6% (N=926/1527) for Havana, 61.2% (N=835/1320) for Santo Domingo, and 54.3% (N=661/1217) for San Juan.[2]

[1] See Chapter 4.3.3.

[2] Without *hay/hayn* the rate of *haber* pluralization documented for San Juan is still significantly higher than the 44% (N=83/190) of pluralized forms obtained by Brown & Rivas (2012: 329) for Caguas, Cayey, and San Juan. These fluctuations could be due to the fact that Brown & Rivas (2012) base their conclusions on a smaller number of observations (total N=190) derived exclusively from semi-directed interviews.

Table 6.2: Significant linguistic factors in the pluralization of presentational *haber* in Havana, Santo Domingo, and San Juan: Numbers, percentages, and weights for the pluralized presentational *haber* construction

Factor groups	*Havana*[a]				*Santo Domingo*[b]				*San Juan*[c]			
	N	*%*	*W(L)*	*W(S)*	*N*	*%*	*W(L)*	*W(S)*	*N*	*%*	*W(L)*	*W(S)*
Verb tense												
All others	819/1298	63.1	.79	.76	720/1103	65.3	.80	.76	622/1014	61.3	.82	.81
Synthetic expressions in present or preterit tense	115/795	14.5	.21	.24	140/739	18.9	.20	.24	62/641	9.7	.18	.19
Range			58	52			60	52			64	62
Production-to-production priming												
Pluralized presentational *haber* construction	556/817	68.1	.68	.67	484/711	68.1	.70	.68	352/558	63.1	.66	.65
First occurrence/distance 20+ clauses	83/297	27.9	.44	.42	123/337	36.5	.47	.44	88/246	35.8	.46	.45
Singular presentational *haber* construction	295/979	30.1	.37	.41	253/794	31.9	.33	.37	244/851	28.7	.38	.39
Range			31	26			37	31			28	26
Comprehension-to-production priming												
Pluralized presentational *haber* construction	113/239	47.3	.62	.58	151/264	57.2	.61	.62	92/175	52.6	.63	.64
Singular presentational *haber* construction	73/204	35.8	.47	.49	63/185	34.1	.47	.46	30/125	24.0	.44	.43
First occurrence/distance 20+ clauses	748/1650	45.3	.41	.43	646/1393	46.4	.42	.41	562/1355	41.5	.44	.43
Range			21	15			19	21			19	21
Typical action-chain position of the noun's referent												
Heads	467/925	50.5	.57	.56	439/815	53.9	.61	.61	338/730	46.3	.58	.58
Tails and settings	467/1168	40.0	.44	.44	421/1027	41.0	.39	.39	346/925	37.4	.42	.42
Range			13	12			22	22			16	16
Absence/presence of negation												
Absent	Not significant				Not significant				559/1225	45.6	.57	.57
Present									125/430	29.1	.43	.43
Range											14	14

Notes: *W* means 'factor weight'. *(L)* means 'model with the lemmas'. *(S)* means 'model with the individual speakers'. The cognitive and social factors were included in the same regression model, but space inhibits tabulating all the results on the same page. The social factors will be presented in **Table 6.12** (see page 149). [a] With the speakers: deviance: 2049.45; AIC: 2075.45; centered input probability: .33; with the lemmas: deviance: 2019.93; AIC: 2045.93; centered input probability: .36. [b] With the speakers: deviance: 1889.31; AIC: 1909.31; centered input probability: .46; with the lemmas: deviance: 1829.93; AIC: 1851.93; centered input probability: .45. [c] With the speakers: deviance: 1549.19; AIC: 1579.19; centered input probability: .32; with the lemmas: deviance: 1519.56; AIC: 1549.56; centered input probability: .35 (AIC=Akaike's Information Criterion)

2. Cognitive factors

In this section, the results of Table 6.2 will be reviewed and it will be detailed how these linguistic constraints relate to the three general cognitive factors introduced in Chapter 3.3.2. Specifically, section 2.1 focuses on the effects of markedness of coding, whereas section 2.2 deals with the way statistical preemption constrains *haber* pluralization. Section 2.3, in turn is concerned with structural priming effects. Section 2.4 investigates the interaction between these cognitive factors. Before turning to the social factors that are considered in this study, section 2.5 presents and compares the constraint rankings for the three speech communities.

2.1 Markedness of coding

This section explores how markedness of coding constrains *haber* pluralization. To this end, section 2.1.1 evaluates which features define the NP of *haber* as a less or a more prototypical subject. Then, section 2.1.2 focuses on the way the absence/presence of negation contributes to portraying the nominal argument of presentational *haber* as an object or as a subject.

2.1.1 The properties of the NP

In Chapter 3.3.2.1, hypothesis 1 claims that the preference for unmarked coding causes speakers to use the pluralized presentational *haber* construction more often with nominal arguments that are more prototypical subjects. However, this hypothesis raises the question as to which features can portray the NP of presentational *haber* as a more prototypical object or subject. In this regard, the typological literature suggests that prototypical subjects refer to agents in events (Comrie, 1989: 66; Dixon, 1979: 86; Du Bois, 1987: 829; Keenan, 1976: 321). However, as we have seen in the Chapter 5.2.1, the nominal argument of presentational *haber* is clearly not an agent, because it is merely present in a stative situation. Still, it is inarguably the case that some entities (say, e.g., *driver*) are intrinsically more likely than others (say, e.g., *invitee*) to fulfill this role. Therefore, with constructions such as presentational *haber*, which do not explicitly construe the nominal argument as an agent or as a patient, entities like *driver* may be perceived as more potential agents, and, thus, as more prototypical subjects (Langacker, 1991: 294) than entities like *invitee.*

In cognitive linguistics, the semantic roles ‘agent’ and ‘patient’ are defined in relation to what Langacker (1991: 283-285) calls the ‘canonical event model’ or the ‘action-chain model’: the head initiates physical activity, resulting “through physical contact, in the transfer of energy to an external object” (Langacker, 1991: 285) and an internal change of state of that entity, the tail of the chain. The semantic roles of agent and patient, in turn, are defined as, respectively, ‘action-chain head’ and ‘action-chain tail’. Additionally, events take place in a particular setting, such that the event model minimally includes three elements: action-chain head/agent, action-chain tail/patient,

and setting. Therefore, to test the first hypothesis, I coded the data for the typical action-chain position of the entity indicated by the isolated noun, for which I relied on the answers to the question in (1).

(1) *Is the referent of the noun highly likely to cause an internal change of state to a second entity without being affected by a third entity first?*
Yes: Typical action-chain head.[3]
e.g., *temblor* 'earth quake', *madre* 'mother', *carro* 'car'…
No: Typical action-chain setting or tail.[4]
e.g., *actividad* 'activity', *víctima* 'victim', *daño* 'damage'…

As predicted by hypothesis 1, Table 6.2 indicates that speakers of the three varieties are more likely to use the pluralized presentational *haber* construction with nouns that refer to typical action-chain heads (shown in example 2), whereas with nouns that refer to typical tails, as in example (3), or settings (see example 4), they prefer the singular presentational *haber* construction.

(2) Humans such as *madre* 'mother', natural phenomena such as *huracán* 'hurricane', self-propelling objects such as *carro* 'car', *tiro* 'gun shot'.

(3) Tangible objects such as *libro* 'book', animate beings that undergo an action such as, for example, *víctima* 'victim' and *invitado* 'invitee'.

(4) *Lugar* 'place', *año* 'year', nominalized events such as *actividad* 'activity', *discusión* 'discussion'.

In contrast, as noted in Chapter 2.3, previous studies on *haber* pluralization have claimed that human vs. non-human reference (e.g., Bentivoglio & Sedano, 1989, 2011: 172-174), the noun's proportion of subject use (Brown & Rivas, 2012), or its temporal boundedness (Rivas & Brown, 2012) are the relevant constraints related to the NP. More particularly, Bentivoglio & Sedano (1989), among many others, have shown that human-reference NPs favor *haber* pluralization. Brown & Rivas (2012), for their part, indicate that *haber* pluralization occurs more often with nouns that are frequently used as subjects in Spanish.

However, I would like to argue that these results actually reflect differences in typical action-chain position. In this regard, although the typological literature shows that animate-reference nouns are indeed more likely to be used as subjects (Ashby & Bentivoglio, 1993; Croft, 2003:130; Du Bois, 1987: 829), it should also be observed that animate reference is only connected to subjecthood through the tendency for animate beings to be agents in events (Dixon, 1979: 86; Du Bois, 1987: 829). In other words, most animate-reference nouns are also typical action-chain heads. Similarly,

[3] Or, in other words, more potential agent.
[4] Or, in other words, more potential setting or patient.

chances are high that nouns of high proportion of subject use also refer to more typical agents.

In order to gain more insight in this matter, I ran two additional, parallel models on the San Juan database, besides the one presented in Table 6.2. In one model, typical action-chain position had been replaced with animacy (animate vs. inanimate reference). In the other, typical action-chain position had been replaced with Brown & Rivas's (2012) factor group 'proportion of noun use as subject' (low to mid vs. high). Let us consider briefly the way these factor groups were operationalized.

Of the two, animacy is perhaps the most self-explanatory. However, it should be added that the data were coded for the animacy of the NP as perceived in the specific context. This way, for instance, *entidades* 'entities' in example (5) was included in the animate category.

(5) Interviewer: ¿Y piensas que el, que el, que el dialecto de San Juan suena mejor?
Participant: Es que no puedo decir "de San Juan" porque *hay* muchas entidades de, de, de, de esas áreas que están acá. Pero quizás… Y tampoco puedo decir que, que las personas están mejor educadas en San Juan, porque hay muchas personas de muy buenas familias que viven en, alre, en la Isla, que también tienen muy buena educación (SJ02M12/SJ200).
Interviewer: 'And do you think that the dialect of San Juan sounds better?'
Participant: 'It's that I cannot say "of San Juan", because *there are*$_{Sing}$ a lot of entities of, of, of, of those areas that are here. But perhaps… And neither can I say that, that the people are better educated in San Juan, because there are a lot of people of very good families that live in, arou, on the Island, that have a very good education as well.'

In turn, elaborating on Brown & Rivas's (2012) methodology, the proportion of subject use of the 172 types of lemmas that appear in the San Juan corpus was established by tracing their grammatical distributions in random 200-clause samples drawn from Davies's (2002-) *Corpus del Español*.[5] Since lemmas with a proportion of subject use above 25% had a similar effect on *haber* pluralization, this factor group was operationalized as shown in (6).

(6) • Used as subject in 0-25% of the cases
• Used as subject in more than 25.5% of the cases

[5] I performed lemma searches in the spoken sections of Davies's (2002-) *Corpus del Español*. If this did not yield enough tokens, the searches were expanded to the entire twentieth-century section of the corpus. When a search produces many results, the corpus offers the possibility to retain only a 200-item random sample of the search results for analysis. In contrast, Brown & Rivas (2012) analyze all search results the corpus provides. However, as my dataset contains much more types of nouns, this proved unfeasible. With my slightly modified method, the same results can be achieved with less coding time per noun type.

Table 6.3: Results for animacy and proportion of noun use as subject in San Juan (parallel models): Numbers, percentages, and weights for the pluralized presentational *haber* construction

Factor groups	*San Juan*			
	N	*%*	*W(L)*	*W(S)*
Animacy[a]				
Animate	308/692	44.5	.58	.56
Inanimate	376/963	39.0	.42	.44
Range			16	12
Proportion of noun use as subject[b]				
Used as subject in more than 25.5% of the 200-clause sample	165/321	51.4	.58	.55
Used as subject in 25% or less of the 200-clause sample	519/1334	38.9	.42	.46
Range			16	9

Notes: *W* means 'factor weight'. *(L)* means 'model with the lemmas'. *(S)* means 'model with the individual speakers'. [a] With the speakers: deviance: 1557.99; AIC: 1587.99; centered input probability: .32; with the lemmas: deviance: 1520.11; AIC: 1550.11; centered input probability: .35. [b] With the speakers: deviance: 1567.73; AIC: 1597.73; centered input probability: .33; with the lemmas: deviance: 1524.26; AIC: 1554.26; centered input probability: .38 (AIC = Akaike's Information Criterion)

Although Table 6.3 shows that, in the San Juan dataset, pluralized *haber* occurs more often with animate-reference NPs and nouns that are frequently used as subjects, this does not necessarily mean that animacy (e.g., Bentivoglio & Sedano, 1989) or the general proportion of subject use of a noun (Brown & Rivas, 2012) condition *haber* pluralization. Rather, as I anticipated, a close comparison of the lemmas of the San Juan corpus (see Table 6.4) unveils that most animate-reference nouns and most nouns that are frequently used as subjects also refer to typical action-chain heads. This is reflected in Table 6.5, which shows that the model with typical action-chain position has significantly lower deviance, Akaike's Information Criterion, and Bayesian Information Criterion scores than the other two, suggesting a better fit.[6]

[6] Model fits can easily be contrasted in R (R Core Team, 2013). For each of the three models, I stored Rbrul's (Johnson, 2014) step-down (called *down.model*) in a variable. Afterwards, I instructed R to evaluate the fit of the three models, applying the function *anova()* to the model objects. This function outputs the three measures of model fit displayed in Table 6.5 and performs likelihood-ratio tests to establish the significance level of the differences between them (Baayen, 2008: 183). It should also be observed that, as three models with an equal number of degrees of freedom are compared in this specific case (i.e., one binary factor group is substituted for another one), even the slightest reduction in deviance/Akaike's Information Criterion/Bayesian Information Criterion is highly significant (Tagliamonte, 2006: 148-149).

Table 6.4: Lemmas of the nouns that occur in the San Juan corpus, by proportion of subject use in 200-token samples drawn from Davies (2002-)

Low to mid	*%*	*Low to mid*	*%*	*Low to mid*	*%*	*Low to mid*	*%*	*Low to mid*	*%*	*Low to mid*	*%*	*High*	*%*	*High*	*%*
cuido	0.0	salón	6.5	evolución	10.5	cambio	13.0	tos	17.0	pez	22.0	invitado	26.0	motorista	43.5
lo	0.0	tanto	6.5	tienda	10.5	carretera	13.0	modalidad	17.5	traductor	22.0	persona	26.0	madre	44.0
año	0.5	contacto	7.5	caña	11.0	parque	13.0	pelea	17.5	corporación	22.5	vecino	26.5	María	44.5
vez	0.5	discusión	7.5	daño	11.0	tradición	13.0	gente	18.0	crítica	22.5	deambulante	27.0	temblor	44.5
día	1.5	madera	7.5	fiesta	11.0	vianda	13.0	inundación	18.0	gran sociedad	22.5	pájaro	27.5	señor	51.5
mes	2.0	maltrato	7.5	muerte	11.0	gallina	13.5	problema	18.0	maremoto	22.5	monja	28.0	papá	62.5
semana	3.0	pantalón	7.5	pastel	11.0	amistad	14.0	posibilidad	18.5	variación	23.0	estudiante	29.4		
sillón	3.0	parranda	7.5	piña	11.0	huevo	14.0	sentimiento	18.5	factor	23.5	niño	30.0		
butaca	3.5	hueso	8.0	sector	11.0	regla	14.0	comida	19.0	nieto	23.5	tormenta	30.0		
casa	3.5	negocio	8.0	animal	11.5	árbol	14.5	lobo	19.0	paloma	23.5	asiático	30.5		
corral	3.5	baño	8.5	asalto	11.5	carro	15.0	monstruo	19.0	grupo de	24.0	ladrón	30.5		
momento	3.5	fruta	8.5	carnaval	11.5	grosella	15.0	recuerdo	19.0	indígena	24.5	compañero	31.0		
peso	3.5	reunión	8.5	conocimiento	11.5	víctima	15.0	ley	19.5	terremoto	25.0	ardilla	31.5		
sitio	4.0	apartamento	9.0	pescado	11.5	iglesia	15.5	león	20.0			huracán	32.0		
cuarto	4.5	clase	9.0	almuerzo	12.0	matrimonio	15.5	medio de comunicación	20.0			joven	32.0		
restaurante	4.5	condominio	9.0	beneficio	12.0	pala	15.5	sonido	20.0			alumno	34.5		
porche	5.0	juego	9.0	curso	12.0	pino	15.5	indicio	20.5			vegetariano	34.5		
lugar	5.5	juey	9.0	llave	12.0	tipo	15.5	diferencia	21.0			maestro	35.0		
paso de río	5.5	castigo	9.5	vecindario	12.0	truco	15.5	hijo	21.0			corpulento	35.6		
uniforme	5.5	escuela	9.5	bandera	12.5	libro	16.0	influencia	21.0			tonto	35.6		
mango	6.0	faceta	9.5	cosa	12.5	muerto	16.0	edificación	21.5			pelotero	38.5		
actividad	6.5	plato	9.5	experiencia	12.5	frase	16.5	habitante	21.5			padre	39.0		
entidad	6.5	china	9.8	piso	12.5	palabra	16.5	sapo	21.5			mujer	39.5		
nivel	6.5	ciudad	10.5	talento	12.5	peligro	16.5	situación	21.5			hermana	40.0		
quiosco	6.5	estilo	10.5	café	13.0	visita	16.5	familia	22.0			muchacho	41.5		

Table 6.5: Comparison of model fit indicators for the three models for the San Juan dataset

Models	*Akaike's Information Criterion*		*Bayesian Information Criterion*		*Deviance*	
	Nouns	Speakers	Nouns	Speakers	Nouns	Speakers
Model with typical action-chain position	1549.6	1579.2	1630.7	1660.4	1519.6	1549.2
Model with animacy	1550.1	1588.0	1631.3	1669.2	1520.1	1558.0
Model with proportion of noun use as subject	1554.3	1597.7	1635.4	1678.9	1524.3	1567.7

Notes: Likelihood-ratio tests show that p = 0.001 for the differences in model fit. Lower values are suggestive of better model fits

The results presented so far support three intermediate conclusions. First, the qualitative and quantitative evidence distilled from the Puerto Rican dataset suggests that the typical action-chain position of the noun's referent is a better predictor for speakers' behavior than, respectively, animacy or a noun's general proportion of subject use.[7] Second, as claimed by hypothesis 1, the data in Table 6.2 show that speakers of the three varieties tend to encode entities that are likely to be agents in events as subjects.[8] Third, in lieu of supporting Rivas & Brown's (2012: 87) claims that temporal persistence (in terms of independent existence and reference) is a feature of prototypical subjects and that 'stage-level nouns'[9] disfavor *haber* pluralization because they are not temporally persistent, the results of this study suggest that these authors' findings reflect differences in typical action-chain position, as stage-level nouns (e.g., *años* 'years' in example 7, *actividades* 'activities' in example 8, *peleas* 'fights' in example 9) refer to typical settings of action chains rather than to typical heads.

(7) Pero después de ahí *hubo* <u>años</u> que no apareció un regalo tampoco (SD20H12/RD2679).
'But afterwards, *there were*$_{Sing}$ <u>years</u> that there didn't appear a gift either.'

[7] Of course, in principle, nothing prohibits that a noun's typical action-chain position is expressed mentally as a 'grammatical relation probability', which, in turn, conditions the variation, as Brown & Rivas (2012) argue. Yet, the results seem to argue against this idea, because if the variation were conditioned by such a probability, we would expect subjects to decide on an item-by-item basis whether to encode the NP argument as a subject or not, rather than depend on the generalization (Goldberg, 2006a:94-102; Lakoff, 1987:147) that agent-like nouns tend to be used as subjects. In statistical terms, such item-by-item decisions would imply that the influence of the grouping factor (in this case, the individual lemmas) would rise above that of the fixed factor typical action-chain position of the noun's referent and, hence, that Rbrul would remove this independent variable from the model (Johnson, 2009:365). This is not the case.

[8] In other words, the results support that more prototypical subjects are preferentially encoded as subjects.

[9] For example, event nouns, deverbal nouns, and temporal nouns (Rivas & Brown, 2012: 81; see Chapter 2.2.6 for discussion).

(8) E, *hay* muchas actividades, viernes, sábado y domingo son actividades (SJ16H21/SJ1937).
'Er, *there are*$_{Sing}$ a lot of activities, Friday, Saturday, and Sunday are activities.'

(9) Eso, *hubo* muchas peleas, e, mujeres y mujeres y hombres y hombres y, y de todo, o sea, a, eso se peleó mucho ahí (SJ04M22/SJ473).
'Er, *there were*$_{Sing}$ a lot of fights, er, women and women and men and men, and, and, a bit of everything, that is, a, that, they fought a lot over there.'

2.1.2 The absence/presence of negation

In Chapter 2.2 it was said that two of the more recent investigations of *haber* pluralization (D'Aquino-Ruiz, 2004; Quintanilla-Aguilar, 2009) have examined the effect of the absence/presence of negation, with comparable results: the presence of negation disfavors the pluralized presentational *haber* construction. Similarly, Table 6.2 shows that the presence of negation disfavors the pluralized presentational *haber* construction in San Juan, which is also evident from contrastive examples such as (10). In Havana and Santo Domingo, however, this factor group has no influence whatsoever.

(10) *Habían* menos, no *había* tantos salones (SJ07H21/SJ886).
'*There were*$_{Plur}$ less, *there weren't*$_{Sing}$ as many class rooms.'

Although the absence/presence of negation has not been linked to any specific grammatical function in the typological literature, I would like to argue here that the influence of this factor group constitutes an additional reflex in this variation of the preference for unmarked coding. Specifically, recall that in Chapter 5.2.1 it was suggested that the POINTING-OUT ICM implies that the NP is always interpreted as referring to a specific instance rather than to a type (Prince, 1992: 299-300). In negative clauses, however, the reference of the nominal argument becomes suspended (Brown & Rivas, 2012: 327; Keenan, 1976: 318; Suñer, 1982: 85). As a consequence, it is interpreted as nonspecific indefinite, that is, as being "identifiable only as a type, not as a specific instance or token" (Croft, 2003: 132).

In this regard, the typological literature supports that nonspecific indefinite is the prototypical definiteness/specificity value of direct objects (Croft, 2003: 132; Du Bois, 1987: 847; Keenan, 1976: 319), for which the San Juan results can be interpreted as an effect of the preference for unmarked coding. Yet, this does not account for the Havana and Santo Domingo data. However, as these two communities also display slightly higher overall rates of *haber* pluralization, the results may indicate that, in Havana and Santo Domingo, the pluralized presentational *haber* construction has invaded the non-specific indefinite conceptual territory. This, in turn, could suggest an ongoing linguistic change (Company-Company, 2003: 26).

2.2 Statistical preemption

Earlier variationist studies have almost consistently documented *haber* pluralization to occur far more frequently in the imperfect tense than in the preterit and present tenses.[10] In Chapter 3.3.2.2, Hypotheses 2a-c propose that the effect of the verb tense is the reflex in this variation of statistical preemption, that is, the tendency to use a partially lexically filled instance of a construction rather than constructing a novel expression based on a more abstract construction schema when both the abstract schema and the entrenched instance could encode the conceptualization equally well. Particularly, it was argued that for the tenses that were mainly used in presentational *haber* expressions before the actuation of the change,[11] a partially lexically filled instance of <**AdvP** *haber* **Obj**>, conserved through repetition, preempts the use of the pluralized presentational *haber* construction for expressions that do not involve aspectual or modal auxiliaries.

These hypotheses raise two questions: first, when did the variation that affects presentational *haber* emerge as a community-wide phenomenon and, second, which forms of the verb enjoyed a relatively high frequency in a variety of constructions before this happened? The answer to the first question can only be tentative, as it is difficult to know when and how the variation that affects presentational *haber* started exactly in Caribbean Spanish.[12] For Buenos Aires, Fontanella de Weinberg (1992b: 39) has shown that the alternations between pluralized and singular *haber* already occur with some frequency in written discourse from the eighteenth century onward. Since there is usually a considerable lag between the actuation of a linguistic change and its trickling down into writing, the variation probably emerged somewhere in the seventeenth century. As we will see in Chapter 7.2, this blends in nicely with research in historical linguistics showing that the most prominent features of American Spanish stem from a koine variety[13] (De Granda, 1994: Chap. 1; Fontanella de Weinberg, 1992a: Chap. 1) that emerged through language and dialect contact during that century.

Regarding the answer to the second question, judging from present-day Spanish, present-tense *hay* was probably used predominantly in presentational *haber* expressions before <**AdvP** *haber* **Subj**> emerged as a conventional alternative, because the verb developed a special form for the simple present in its presentational use (Hernández-Díaz, 2006: 1141-1145). Therefore, it is to be expected that this tense will resist the pluralized presentational *haber* construction most. Since the preterit perfect (e.g., *hubo hablado* 'had spoken') has hardly ever been used in Spanish (Bull,

[10] See Chapter 2.2.

[11] See Chapter 3.1.8.

[12] See Chapter 7.2 for a tentative proposal in this regard.

[13] See, for example, Kerswill & Williams (2000) and Trudgill (1986) for discussion of koineization.

Cantón, Cord, Farley, Finan, Jacobs, Jaeger, Koons, & Tuegel, 1947: 456; Hills & Anderson, 1929: 604; Keniston, 1937: 193; Rivas & Brown, 2013: 115), preterit *hubo* constitutes another likely candidate for entrenchment in the singular presentational *haber* construction.

In Table 6.6, the frequencies of the third-person singular forms of *haber* in a sixteenth-century Latin American subset of the Royal Spanish Academy's *Corpus Diacrónico del Español* (Real Academia Española, 2008a-) are tabulated. As I suspected, the corpus searches show that before presentational *haber* was subject to large-scale variation in Latin American Spanish, its present- and preterit-tense forms occurred primarily in presentational clauses. This suggests that the most salient representations of these verb forms were <**AdvP** *hay* **Obj**> and <**AdvP** *hubo* **Obj**>. The other tense forms, on the other hand, were either used more productively (spread over more different constructions) or are restricted to a very low frequency in the corpus (N < 100), which indicates that their independent forms probably also constituted their strongest cognitive representations. In other words, this distribution suggests two relevant types of presentational *haber* expressions, listed in (11).

(11) • Synthetic expressions[14] in the present and preterit tense
• All other expressions[15]

Turning now to the results for this factor group, Table 6.2 shows that in the three varieties, statistical preemption constitutes the strongest overall cognitive constraint on the variation.[16] Additionally, as predicted by hypothesis 2a-b, the pluralized presentational *haber* construction is unlikely to be used with the present or preterit (for which Table 6.6 suggests an entrenched instance of <**AdvP** *haber* **Obj**>) when the coding of the conceptual import does not call for aspectual or modal auxiliary constructions. This is illustrated in example (12), where the speaker simply points out that, in the past, there were tsunamis in San Juan.

(12) Y aquí *hubo*, este, maremotos (SJ04M22/SJ493).
'And here, er, *there were*$_{Sing}$ tsunamis.'

[14] That is, without aspectual or modal auxiliaries.
[15] Present- and preterit-tense expressions involving aspectual or modal auxiliaries (e.g., *puede haber* 'there can be$_{Sing}$', *deben haber* 'there have to be$_{Plur}$') and the periphrastic future (e.g., *va a haber* 'there are going to be$_{Sing}$') were also included in the latter group.
[16] It has the highest range.

Table 6.6: Distribution of the third-person singular forms of *haber* across constructions in sixteenth-century Latin American texts from Real Academia Española (2008a-)

Constructions	*Había*		*Hubo*		*Habría*		*Habrá*		*Haya*		*Hubiera*		*Hay*		*Ha habido*	
	N	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*
<*haber* past participle>	1806	52.4	38	6.5	6	10.5	41	13.4	202	32.3	110	50	0	0.0	0	0.0
<*haber de* infinitive>	644	18.7	23	4.0	3	5.3	16	5.2	35	5.6	23	10.5	0	0.0	0	0.0
<*haber que* infinitive>	8	0.2	1	0.2	1	1.8	1	0.3	2	0.3	1	0.5	45	1.3	0	0.0
Possessive *haber*	54	1.6	113	19.4	4	7.0	17	5.5	89	14.2	6	2.7	0	0.0	0	0.0
Presentational *haber*	870	25.2	406	69.8	43	75.4	173	56.4	295	47.2	78	35.5	3440	98.7	6	100
<*ha* time expression>	67	1.9	1	0.2	0	0.0	59	19.2	2	0.3	2	0.9	0	0.0	0	0.0
Total	3449	100	582	100	57	100	307	100	625	100	220	100	3485	100	6	100

Notes: The following parameters were used for the collection of the instances of *haber*: 1492–1600, *Lírica, Narrativa, Breve, Relato breve tradicional*, and *otros*. As initial searches within the Caribbean section of the corpus did not yield enough results, the searches were extended to all of Latin America (Argentina, Bolivia, Chile, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, Puerto Rico, Uruguay, and Venezuela). I did not take administrative and legal documents into account, because these typically contain a very archaic type of language. This results in abundant use of possessive *haber* in the polite imperative mood (*haya*), a usage that had already decayed by that time in other types of sources.

For presentational *haber* expressions that involve other tenses or aspectual or modal auxiliaries, in turn, Table 6.2 shows that the singular presentational *haber* construction is unlikely to be used. This is also evident from contrastive examples such as (13).

(13) Los carnavales es una convocatoria que hace el Estado donde dice: "*Van a haber* carnavales, porque van a pasar unas carrozas y, y *hay* ven, puntos de venta de cerveza." Y la gente va (LH08H12/LH991-LH992).
'The carnivals is a call that the state puts out, in which it says: "*There are going to be*$_{Plur}$ carnivals, because some wagons are going to pass by and, and *there are*$_{Sing}$, outl, beer outlets." And the people attend.'

Additionally, in Chapter 3.3.2.2, hypothesis 2c claimed that complex conceptualizations, involving aspectual or modal auxiliaries (see example 14) would favor the pluralized construction, because these would bypass the entrenched instances of *hay* and *hubo*.

(14) *Pueden haber* expresiones que, que tengan una acepción y una connotación diferentes en el Cibao que las que tienen aquí, en, en, en, en el Sur (SD20H12/RD2706).
'*There can be*$_{Plur}$ expressions that, that may have a different meaning and connotation in the Cibao from those that they have here, in, in, in, in the South.'

Table 6.7 suggests that this is the case, because in expressions with aspectual or modal auxiliaries, pluralized presentational *haber* is used as frequently with the present and the preterit tense as with other tenses, as was already observed in earlier investigations (Hernández-Díaz, 2006:1150; Quintanilla-Aguilar, 2009:164–165).

Table 6.7: Present- and preterit-tense tokens of presentational *haber* in Havana, Santo Domingo, and San Juan, by absence/presence of aspectual or modal auxiliary constructions: Numbers and percentages for the pluralized presentational *haber* construction

Type of expression	*Havana*[a]		*Santo Domingo*[b]		*San Juan*[c]	
	N	*%*	*N*	*%*	*N*	*%*
Presentational *haber* expressions in the present and preterit tense without auxiliary constructions	115/795	14.5	140/739	18.9	62/641	9.7
Presentational *haber* expressions in the present and preterit tense involving auxiliary constructions	133/206	64.6	92/143	64.3	84/124	67.7

Notes: [a] Pearson's Chi-square: 220.32; df: 1; $p < 0.0001$; [b] Pearson's Chi-square: 127.35; df: 1; $p < 0.0001$; [c] Pearson's Chi-square: 226.88; df: 1; $p < 0.0001$

Additionally, although present- and preterit-tense presentational *haber* expressions not involving aspectual or modal auxiliaries were consistently binned together in the tables, this is not to say that both types display similar rates of pluralization. Rather, Table 6.8 indicates that, in synthetic expressions, *haber* pluralization occurs significantly more often with the preterit than with the present tense. This is especially true for Havana and Santo Domingo, where the frequency of <**AdvP** *hubieron* **Subj**> approximates and crosses, respectively, the 50% threshold. This pattern is readily accounted for by statistical preemption: although *hubo* rarely appears outside of presentational *haber* expressions in spontaneous discourse, every native speaker of Spanish will have observed it a limited number of times in four constructions: the singular presentational *haber* construction, the modal constructions <*hubo de* infinitive> '<have to infinitive>' and <*hubo que* infinitive> '<have to infinitive>', and the preterit perfect construction (e.g., *hubo hablado* 'had spoken'). In contrast, in every type of discourse, *hay* only appears in two constructions: the singular presentational *haber* construction and the impersonal obligation modal <*hay que* infinitive> '<have to infinitive>'. Consequently, speakers have more evidence that the preterit of *haber* can occur outside of **<AdvP** *haber* **Obj>** than they have for the present tense and, as a result, the preempting effect that is caused by the latter is stronger than the one that goes out from the former.

For the preterit, Table 6.8 also shows that, as was the case for the absence/presence of negation, the division of labor between the pluralized and the singular presentational *haber* construction seems to be stricter in San Juan than in Havana and Santo Domingo. That is to say, whereas in San Juan <**AdvP** *hubo* **Obj**> is still the dominant variant to construe preterit presentational *haber* expressions that do not involve aspectual or modal auxiliaries, the pluralized variant is already well on its way to take over this function in Havana. Moreover, in Santo Domingo, the frequency of <**AdvP** *hubieron* **Subj**> already crosses the 50% threshold. As noted for the absence/presence of negation, the apparent loosening of the restrictions on the use of pluralized presentational *haber* expressions in Havana and Santo Domingo vis-à-vis San Juan may suggest an ongoing linguistic change.

Table 6.8: Present- and preterit-tense tokens of presentational *haber* without aspectual or modal auxiliary constructions in Havana, Santo Domingo, and San Juan: Numbers and percentages for the pluralized presentational *haber* construction

Tense	*Havana*[a]		*Santo Domingo*[b]		*San Juan*[c]	
	N	*%*	*N*	*%*	*N*	*%*
Present	8/566	1.4	24/521	4.6	21/433	4.8
Preterit	107/229	46.7	116/218	53.2	41/208	19.7

Notes: [a] Yates's Chi-square: 266.89; df: 1; $p < 0.0001$; [b] Pearson's Chi-square: 236.45; df: 1; $p < 0.0001$; [c] Pearson's Chi-square: 35.52; df:1; $p < 0.0001$

Finally, contrary to the analysis that was developed in this section, Waltereit & Detges (2008: 27) state that

> [i]n spoken language, presentational constructions are most frequently used in the present tense. Hence, of all the forms of *haber* + NP, the irregular present tense *hay* + NP is the most solidly entrenched one. Reanalyses based on low frequency will therefore more likely occur in non-present tenses. Bentivoglio & Sedano (1989) show that the tense most affected by this reanalysis is indeed the imperfect, which is the least frequent of the two Spanish past tenses.

However, without positing a competition between two construction schemas, entrenched instances, and differing strengths of statistical preemption, it is difficult to explain why the pluralized presentational *haber* construction is used less with the preterit vis-à-vis other non-present tenses. In addition, Waltereit & Detges's (2008: 27) analysis sets off on the wrong premises. That is, although it is true that the imperfect is less frequent than the preterit tense when we consider all Spanish verbs (Bull et al., 1947: 456; Hills & Anderson, 1929: 604), this is not the case for presentational *haber* expressions, which occur most often with the simple present and imperfect tense (Brown & Rivas, 2012: 79, 2013: 115). In addition, if the effect of the verb tense were somehow due to differing degrees of morphophonological contrast between the pluralized and singular forms of *haber* across tenses (e.g., Bentivoglio & Sedano, 2011: 174; Hernández-Díaz, 2006: 1151) rather than to statistical preemption, then we would not expect to find that auxiliary constructions, which display the same contrasts, favor the pluralized presentational *haber* construction. In other words, the true constraint imposed by the verb tense seems to be statistical preemption, that is, that

> more specific items are preferentially produced over items that are licensed but are represented more abstractly, as long as the items share the same semantic and pragmatic constraints (Goldberg, 2006a: 94).

2.3 Structural priming

In Chapter 3.3.2.3, hypothesis 3 suggests that speakers' preference for one of the alternatives of the presentational construction with *haber* would be influenced by the variant they have used or processed earlier, if any. This prediction is based upon the long-standing observation in psycholinguistics that language users tend to pick up and recycle (unintentionally and unconsciously) construction patterns they have (heard) used before, without necessarily repeating the specific words that appear in these structures.[17] In the psycholinguistic literature, this is called 'structural priming'.[18]

[17] However, priming effects are stronger when the same verb is repeated in both the prime and the target clause (Bock & Griffin, 2000: 188; Pickering & Branigan, 1998: 640). Still, the fact that tense, aspect, mood, and number variations do not affect the magnitude of the priming effect (Pickering & Branigan, 1998: 640-645) suggests that it cannot be the case that speakers simply repeat (parts of) the expression they have processed (Bock & Griffin, 2000: 188; Pickering & Branigan, 1998: 646). Rather, it seems that there is something going on at a more abstract level. See Pickering & Ferreira (2008) for a synthesis.

[18] This phenomenon is also sometimes referred to as 'syntactic priming' or 'syntactic persistence'. However, according to Pickering & Ferreira (2008: 427-428), 'structural priming' is a more adequate term, because, in

Although usually couched in different terms and investigated using different research paradigms, the same tendency has been documented in variationist sociolinguistics (e.g., Labov, 1994: 550-566; Travis, 2005: 340, 2007: 115; Weiner & Labov, 1983: 52), where structural priming has often been shown to constitute one of the most important constraints on morphosyntactic variation (Labov, 1994: 550-566; Martín-Butragueño, 1999: 232; Pereira-Scherre & Naro, 1992: 8; Weiner & Labov, 1983: 53).

Psycholinguistic experiments have shown priming effects to last for at least ten intervening clauses (Bock & Griffin, 2000: 186; Bock, Dell, Chang, & Onishi, 2007: 452; Pickering & Ferreira, 2008:447) and to be modality-independent (Bock et al., 2007: 454; Pickering & Ferreira, 2008: 440-441). Therefore, the data were coded for the type of last token that was provided by the interviewer (comprehension-to-production priming) and the speaker (production-to-production priming) and the number of conjugated verbs that occur between these tokens and the case at hand. While coding, the occurrences were binned together in five-clause lag groups[19] up until reaching a 20-clause lag and the occurrences in which speakers repeated the verb form and the presentational *haber* construction were separated from those in which they only repeated the construction. This resulted in a total of seventeen factors for both factor groups. However, as the initial results displayed a similar priming effect up until a twenty-clause lag, independently of whether speakers would repeat the same verb form or not, the factors were collapsed into the broader categories listed in (15).

(15) • First occurrence/distance 20+ clauses
• Primed with the pluralized presentational *haber* construction
• Primed with the singular presentational *haber* construction

Turning now to the results for this cognitive factor, Table 6.2 shows that whenever speakers have used a pluralized presentational *haber* clause, they are more prone to use another one. This is the case whether or not they repeat the same verb form, at least, if the next variable context is situated within a twenty-clause range. The same results were obtained for the singular presentational *haber* construction. Similarly, when speakers have processed a pluralized presentational *haber* clause, they are more likely to utter an expression based on the pluralized construction pattern and vice versa.

At this point, however, it should be observed that Table 6.2 does not exclude the possibility that these results are due to certain speakers having very high or very low baseline rates of *haber* pluralization. If this is the case, the singular or pluralized

---

principle, all levels of linguistic structure can be primed (as opposed to only the syntactic one) and priming does not always involve persistence. It should also be mentioned that there are other types of priming as well, which, however, will not be discussed in this study.

[19] For example, lag: 0-4 clauses; 5-9 clauses, etc.

tokens of these speakers following singular or pluralized instances could lead Rbrul to identify a structural priming effect where there is none. In addition, as *haber* pluralization only occurs rather limitedly in present- and preterit-tense synthetic expressions, it could be the case that these tokens have a similar effect on contexts following a singular presentational *haber* instance. To control for these possible sources of error, first, all synthetic present- and preterit-tense tokens were excluded and, subsequently, the speakers whose rates of pluralization in these restricted datasets were either below 30% or above 75% (Havana: 9 speakers; Santo Domingo: 8 speakers; San Juan: 6 speakers) were removed.

As Table 6.9 shows, when I run the models on this subset of the data, the priming effect remains intact for production-to-production priming, with almost identical factor weights. Since this procedure eliminates about 60% of the tokens (N=3265/5589), for comprehension-to-production priming, the number of observations for the primed conditions drops to the point that no conclusions can be drawn up without running "the risk of taking too little data far too seriously" (Paolillo, 2013: 114). Still, the robustness of the production-to-production priming effect suggests that *haber* pluralization is subject to structural priming. This, in turn, advocates in favor of treating the variation as a competition between two argument-structure constructions, because, without an overarching construction schema that could be repeated, we would not expect plurals to prime plurals and singulars to prime singulars regardless of variations in verb form.

Table 6.9: Production-to-production priming effects for speakers representing robust variability in Havana, Santo Domingo, and San Juan, without present- and preterit-tense synthetic expressions: Numbers, percentages, and weights for the pluralized presentational *haber* construction

Type of last occurrence	*Havana*				*Santo Domingo*				*San Juan*			
	N	*%*	*W(L)*	*W(S)*	*N*	*%*	*W(L)*	*W(S)*	*N*	*%*	*W(L)*	*W(S)*
Pluralized presentational *haber* construction	231/338	68.3	.66	.64	229/314	72.9	.67	.64	207/301	68.8	.63	.64
First occurrence/ distance 20+ clauses	37/87	42.6	.36	.38	74/112	66.1	.45	.50	50/93	53.8	.44	.43
Singular presentational *haber* construction	174/341	51.0	.47	.48	151/343	44.0	.38	.34	180/395	54.4	.43	.43
Range			19	16			29	30			20	21

Notes: *W* means 'factor weight'. *(L)* means 'model with the lemmas'. *(S)* means 'model with the individual speakers'

Finally, priming effects also seem to account for the cases in which the verb agrees with a direct-object pronoun, exemplified in (16).

(16) Interviewer: ¿Este, tú piensas que *pueden haber* diferencias entre las regiones del país en cuanto a comida?
Participant: Bueno, los, t, *tienen que haber<u>las</u>*, porque, por ejemplo, en el Sur se comen más granos (SD19M21/ RD2551).
Interviewer: 'Er, do you think that *there can be*$_{Plur}$ differences between the regions of the country regarding food?'
Participant: [Well, the, t, *there have to be*$_{Plur}$ <u>them</u>$_{Acc}$, because, for example, in the South they eat more grains.]
Participant: 'Well, the, t, *there have to be*$_{Plur}$, because, for example, in the South they eat more grains.'

That is, Table 6.10 (as example 16) shows that the vast majority of the examples of this type occur in contexts primed with <**AdvP** *haber* **Subj**> (Havana: 73.5% N=36/49; Santo Domingo: 82.8% N=24/29; San Juan: 71.0% N=22/31). Hence, rather than constituting strong evidence arguing against the main hypothesis, these results may suggest that priming effects cause individual speakers to reanalyze the direct-object pronoun (a syntactically motivated class of pronouns) as a hearer-new subject pronoun (a pragmatically motivated class of subject pronouns). Still, this appears to be an online phenomenon, because some speakers use this agreement pattern multiple times, whereas others do not use it at all. If the reanalysis of the pronoun were a change in progress, one would expect to find clear social patterning,[20] which is not the case.

Table 6.10: Presentational *haber* tokens that co-occur with object pronouns in Havana, Santo Domingo, and San Juan, by production-to-production priming and comprehension-to-production priming: Numbers and percentages for the pluralized presentational *haber* construction

Type of last occurrence	*Havana*		*Santo Domingo*		*San Juan*	
	N	*%*	*N*	*%*	*N*	*%*
First occurrence/distance 20+ clauses	0/2	0.0	1/2	50.0	0/2	0.0
Singular presentational *haber* construction	13/70	18.6	5/49	10.2	9/47	16.1
Pluralized presentational *haber* construction	36/59	61.0	24/54	44.4	22/48	45.8
Total	49/131	37.4	30/105	28.6	31/106	29.2

Note: Two-tailed Fisher's exact tests show $p < 0.0001$ for the three communities

[20] See Chapter 2.2.1.

2.4 Interaction between the cognitive factors

Up until now, the discussion has been concerned with the way the individual cognitive constraints shape *haber* pluralization when they are considered jointly with the other cognitive factors, the social factors, and the random variation due to individual speakers and lemmas. What has not been considered is the way these factors work in tandem to promote one of the variants or, conversely, interact to cancel each other's effect. As Tagliamonte & Baayen (2012:163-164) observe, disentangling this complex interplay of constraints goes beyond the capabilities of a mixed-effects regression model,[21] but conditional inference trees are very well suited for such a task.[22]

In the conditional inference trees displayed in Figure 6.1-Figure 6.3, only the linguistic factor groups that turned out to be significant in the relevant mixed-effects models were included. Like Table 6.2, the conditional inference trees suggest that statistical preemption constitutes the most important cognitive constraint on the variation, because the verb tense forms the topmost branching node in the three figures. For Havana, the left-hand side of Figure 6.1 (nodes [2] and [4]) also unveils a complex interaction between comprehension-to-production priming, production-to-production priming, and statistical preemption. Particularly, for the 'all others' group of expressions, the first factor is only significant in contexts primed by the speaker with the pluralized construction. This is exemplified in (17), where both the speaker and the interviewer use a pluralized presentational *haber* expression before the speaker utters the last pluralized *haber* clause.

[21] Although mixed-effects regression models can accommodate pairwise interactions between factor groups, they cannot consider the interactions between all factor groups at the same time (Tagliamonte & Baayen, 2012: 164).

[22] As noted in Chapter 4.4.2, another reason to use a combination of these statistical methods is that they rest on completely different distributional assumptions: whereas mixed-effects regression presupposes datasets that are more or less equally distributed over the different factors that are put to the test, conditional inference trees do not require this kind of data structure. This way, if we find similar results with these two complementary methods, we can be more confident that the patterns that emerge from the data are not due to distributional biases.

(17) Interviewer: ¿Este, y durante tu, el tiempo que tú llevas aquí, este, *han habido* muchos cambios aquí?
Participant: ¿Cambios, a, en qué sentido?
Interviewer: ¿No sé, este, que han venido muchas personas nuevas, este, que se ha, bueno, no sé?
Participant: Sí.
Interviewer: O, o…
Participant: Sí.
Interviewer: ¿Que se ha creado algún sitio nuevo o…?
Participant: Bueno,…
Interviewer: ¿Un programa?
Participant: *Han, han habido* m, *han habido* serios cambios. Por la parte social, e, por la parte íntima mía. Amistades, he tenido que hacer amistades nuevas durante muchos años porque casi todos mis amigos han, se han marchado. Unos han ido pa' el Norte, y otros pa' el Sur. Totalmente tengo que hacer amistades nuevas. E, proyectos. ¿Cómo no? En el 2007, e, estuve compartiendo con una delegación también, que, donde *habían* varios estudiantes de diversos países: italianos, franceses, holandeses, que ellos vinieron con un proyecto (LH20H12/LH2699-LH2702).
Interviewer: 'Er, and, during your, during the time that you have been living here, er, *have there been*$_{Plur}$ a lot of changes around here?'
Participant: 'Changes, in, in what sense?'
Interviewer: 'I don't know, er, that there have come a lot of new people, er, that they have, well, I don't know.'
Participant: 'Yes'
Interviewer: 'Or, or…'
Participant: 'Yes'
Interviewer: 'That they have created some new place, or…'
Participant: 'Well…'
Interviewer: 'A program?'
Participant: '*There have*$_{Plur}$*, there have been*$_{Plur}$, m, *there have been*$_{Plur}$ serious changes. For the social part, er, intimately. Friends, I've have had to make new friends during a lot of years, because almost all of my friend have, have left. Some have gone North, others have gone South. Totally, I have to make new friends. Er, projects. Sure. In 2007, er, I was sharing with a delegation as well, where, where, *there were*$_{Plur}$ various students of different countries: Italians, French, Dutch. They came with a project.'

In contrast, the right-hand side of Figure 6.1 displays no such interaction. Moreover, comprehension-to-production priming does not even seem to be a relevant factor for synthetic presentational *haber* expressions in the present and the preterit tense. Rather,

nodes [7], [8], and [11] suggest that production-to-production priming works in tandem with the preference for unmarked coding to promote the pluralized presentational *haber* construction for this type of expressions. This is illustrated by the excerpt from the story-reading task cited in example (18), where the use of the pluralized presentational *haber* clause *hubieron unos ladrones* 'there were$_{Plur}$ some thieves' is favored by both the speaker's earlier use of *hubieron* and the fact that *ladrones* 'thieves' is a typical action-chain head.

(18) Sí que ayer, *hubieron* dos lobos que querían devorarme, anteayer *hubieron* unos ladrones que trataban de matarme y *ha habido* dos veces que yo tenía que brincar un abismo de treinta pies de ancho y todo esto fue muy molesto, pero miedo como tal no tuve" (LH02M12/LH191-LH193).
'But yes, yesterday, *there were*$_{Plur}$ two wolfs that wanted to devour me, the day before yesterday, *there were*$_{Plur}$ some thieves that tried to kill me and *there have been*$_{Sing}$ two times that I had to jump a gap thirty feet wide and all of this was really annoying, but fear as such I didn't have.'

For Santo Domingo, the left-hand side of Figure 6.2 (nodes [2], [3], and [6]) displays a similar pattern of interaction for synthetic present- and preterit-tense *haber*. The right-hand side displays two additional interactions. First, nodes [9] and [10] suggest an interaction between statistical preemption, production-to-production priming, and markedness of coding. Specifically, with non-present, non-preterit expressions or expressions involving auxiliary constructions, the noun's typical action-chain position is only significant in contexts primed by the speaker with pluralized presentational *haber*. Still, the preference for unmarked coding and production-to-production priming appear to work in tandem, because the rates of *haber* pluralization are highest when both line up in favor of the pluralized presentational *haber* construction, as in example (19).

(19) Porque *hubieron* sitios que, que *habían* persecuciones todavía. *Habían* unas gentes muy malas, que el presidente cuando eso era Balaguer (SD03H21/RD346-RD348).
'Because *there were*$_{Plur}$ places where, where *there were*$_{Plur}$ still persecutions. *There were*$_{Plur}$ some very bad people, that the president back then was Balaguer.'

Second, nodes [9] and [13] suggest an interaction between comprehension-to-production priming, production-to-production priming, and statistical preemption. That is, for the 'all others' group of expressions, the first factor is only significant in unprimed contexts or contexts primed by the speaker with singular presentational *haber*. In these cases, comprehension-to-production priming is able to cancel production-to-production priming. This is evident from contexts such as the one

provided in example (20), where the speaker appears to be insensitive to the priming effect that one would expect to go out from her earlier use of *hay*. However, at the same time, the fact that comprehension-to-production priming is only relevant for this restricted subset of the data suggests that this factor has a less profound impact than production-to-production priming.

(20) Interviewer: ¿Y *han habido*, o sea, cuando usted, t, o sea, me podría nombrar cinco cosas que existen hoy y que no *habían* cuando usted era niña? ¿Acá en la ba, en el barrio?
Participant: ¿Cómo así? ¿Cómo así?
Interviewer: ¿Este, como por ejemplo que en, a, edificios que, que poní, que, que pusieron, remodelaciones, e, restaurantes?
Participant: Aja okay, que no *habían* cuando yo era niña. Okay. *Hay* muchas cosas que no *habían* (SD10M21/RD1151-RD1153).
Interviewer: And *have there been*$_{Plur}$, that is, when you, that is, could you name me five things that exist today and that *there weren't*$_{Plur}$ when you were a child, here in the nei, in the neighborhood?
Participant: Like what? Like what?
Interviewer: Er, like, for example, that in, a, buildings that that they pu, that, that they have put, remodeling, er, restaurants?
Participant: Aha, okay, that *there weren't*$_{Plur}$ when I was a child. Okay, *there are*$_{Sing}$ a lot of things that *there weren't*$_{Plur}$.

Turning now to San Juan, the left-hand side of Figure 6.3 (nodes [6], [7], and [10]) shows that both modalities of structural priming and the preference for unmarked coding seem to reinforce each other when statistical preemption does not warn speakers against using pluralized *haber*.[23] For instance, in example (21), the earlier use of the pluralized presentational *haber* construction by the interviewer and the speaker, together with the fact that *muchachos* 'kids' is a typical action-chain head probably tipped the balance in favor of this construction.

[23] That is, with expressions involving aspectual, modal, or temporal auxiliaries or verb tenses other than the present and the preterit.

(21) Interviewer: ¿Y que tú recuerdes, *habían* más padres como los tuyos, los tuyos?
Participant: E, ¿que yo recuerde? Pues en el internado había de todo. *Habían* estudiantes que tenían unos padres que no existían, que las cuidaban las nanas, los cuidaban los… *Habían* unos much, muchachos de mucho dinero (SJ04M22/SJ454-SJ457).
Interviewer: 'And, as far as you remember, *were there*$_{Plur}$ more parents like yours, yours?'
Participant: 'Er, as far as I remember? Well, in the boarding school, there was a bit of everything. *There were*$_{Plur}$ students that had parents that didn't exist, who were looked after by the nannies, they were looked after by… *There were*$_{Plur}$ ki, kids with a lot of money.'

Similarly, nodes [2] and [3] suggest that the tendency to use the pluralized presentational *haber* construction in contexts primed by the speaker with this variant is reinforced by the absence of negation, as in example (22).

(22) Y *habían* de aquí. De Puerto Rico, *habían* dos matrimonios, tres, tres matrimonios y no, no nos conocíamos porque eran de la isla, de por ahí (SJ15M21/SJ1853).
'And *there were*$_{Plur}$ from here. From Puerto Rico *there were*$_{Plur}$ two couples, three, three couples and we didn't, we didn't know each other, because they were from the island, from around there.'

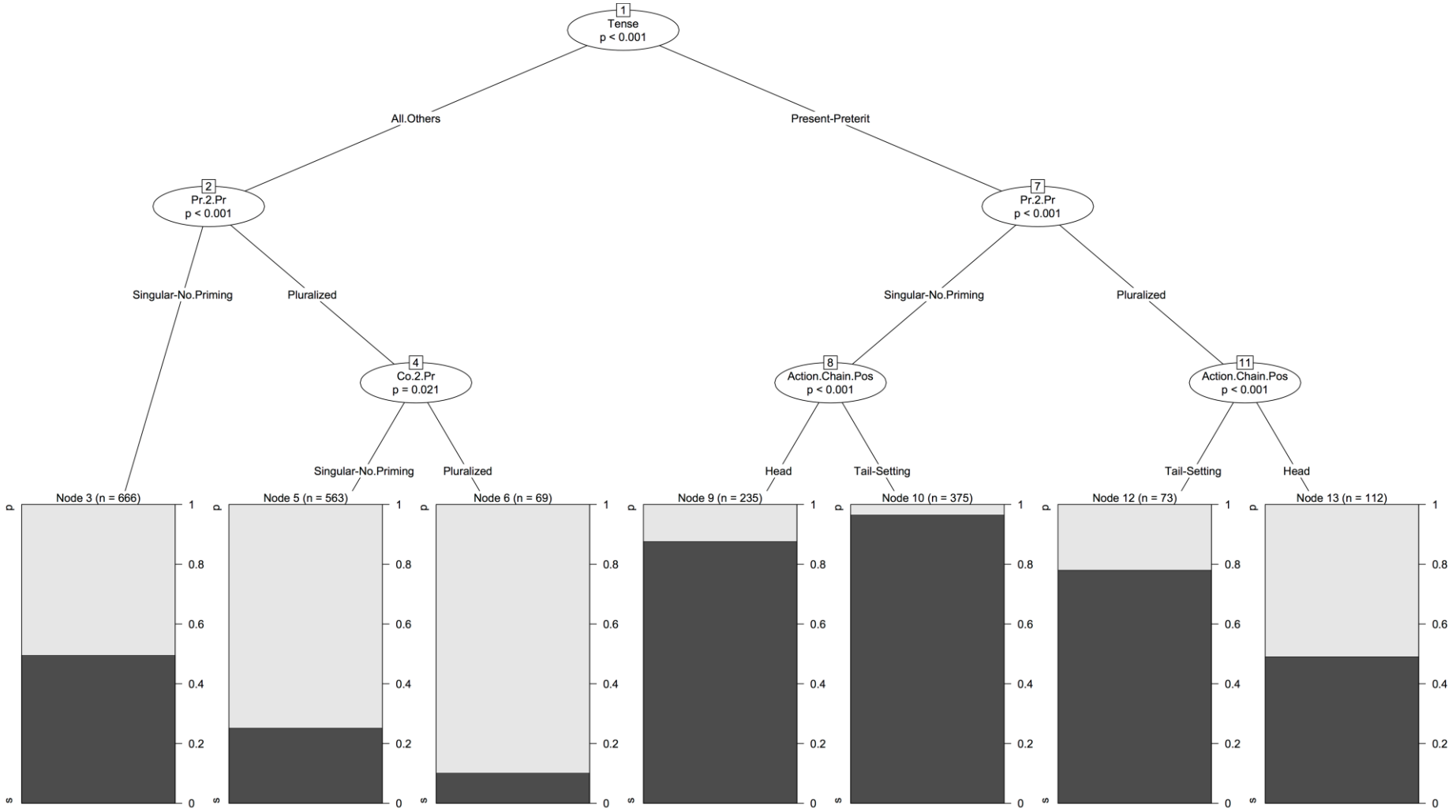
In turn, the right-hand side of Figure 6.3 shows that, for synthetic expressions in the simple present and preterit tense, production-to-production priming appears to operate more independently (see node [13]), because in contexts primed by the speaker with the pluralized presentational *haber* construction, neither the preference for unmarked coding nor comprehension-to-production priming impose significant constraints. In example (23), for instance, the interviewer's earlier mention of *hay* and the fact that *fiesta patronal* 'patron saint celebration' refers to a typical action-chain setting do not cause the speaker to use an expression based on <**AdvP** *haber* **Obj**>. Rather, she continues with the pluralized presentational *haber* construction, which she had already used multiple times before in the immediate context.

(23) Participant: Se pueden comer en todos los momentos, porque, por lo menos en mi casa *hayn* pasteles todos, toda la sem, todo el año. Pero, este, sí, *hayn* platos como que es, específicos de diciembre. Como el arroz con gandules, el lechón, el pastel, las, el arroz con dulce, tembleque.
Interviewer: ¿Y que tú recuerdes siempre ha sido así o *han habido* cambios a este respecto?
Participant: Pues, e, cuando yo era más pequeña se mataba el lechón en casa, mi casa de mi abuela. Se compró todos los, lechones y se mataban allí, y allí los hacían.
Interviewer: ¿Y los asaban?
Participant: Y los, exactamente, ahora no, ahora, pues, ellos los compran hechos.
Interviewer: ¿Y *hay* otras tradiciones por acá, este, fiestas patronales, carnavales?
Participant: Aquí *hayn* fiestas patronales en todos los municipios (SJ05M12/SJ653-SJ657).
Participant: 'They can be eaten at all times, because, at least at my home, *there are*$_{Plur}$ *pasteles* every, all week, all year round. But, er, yes, *there are*$_{Plur}$ dishes that, spe, specific of December. Like rice with pigeon peas, suckling pig, *pastel*, the, rice pudding, *tembleque*.'[24]
Interviewer: 'And as far as you remember, has it always been like that or *have there been*$_{Plur}$ changes in this regard?'
Participant: 'Well, er, when I was smaller, they killed the suckling pig at home, my home of my grandmother. They bought every, suckling pigs and they killed them over there and there they made them.'
Interviewer: 'And you grilled them?'
Participant: 'And them, exactly, not nowadays, nowadays, they buy them ready-made.'
Interviewer: 'And *are there*$_{Sing}$ other traditions around here, er, Patron Saint celebrations, carnivals?'
Participant: 'Here *there are*$_{Plur}$ patron saint celebrations in every town.'

Finally, nodes [13] and [15] suggest an interaction between markedness of coding, production-to-production priming, and statistical preemption. Specifically, for synthetic expressions in the present and preterit tense, the absence/presence of negation is only a relevant constraint in unprimed contexts or contexts primed by the speaker with singular presentational *haber*. In these cases, the absence of negation attenuates the tendency to use the singular presentational *haber* construction.

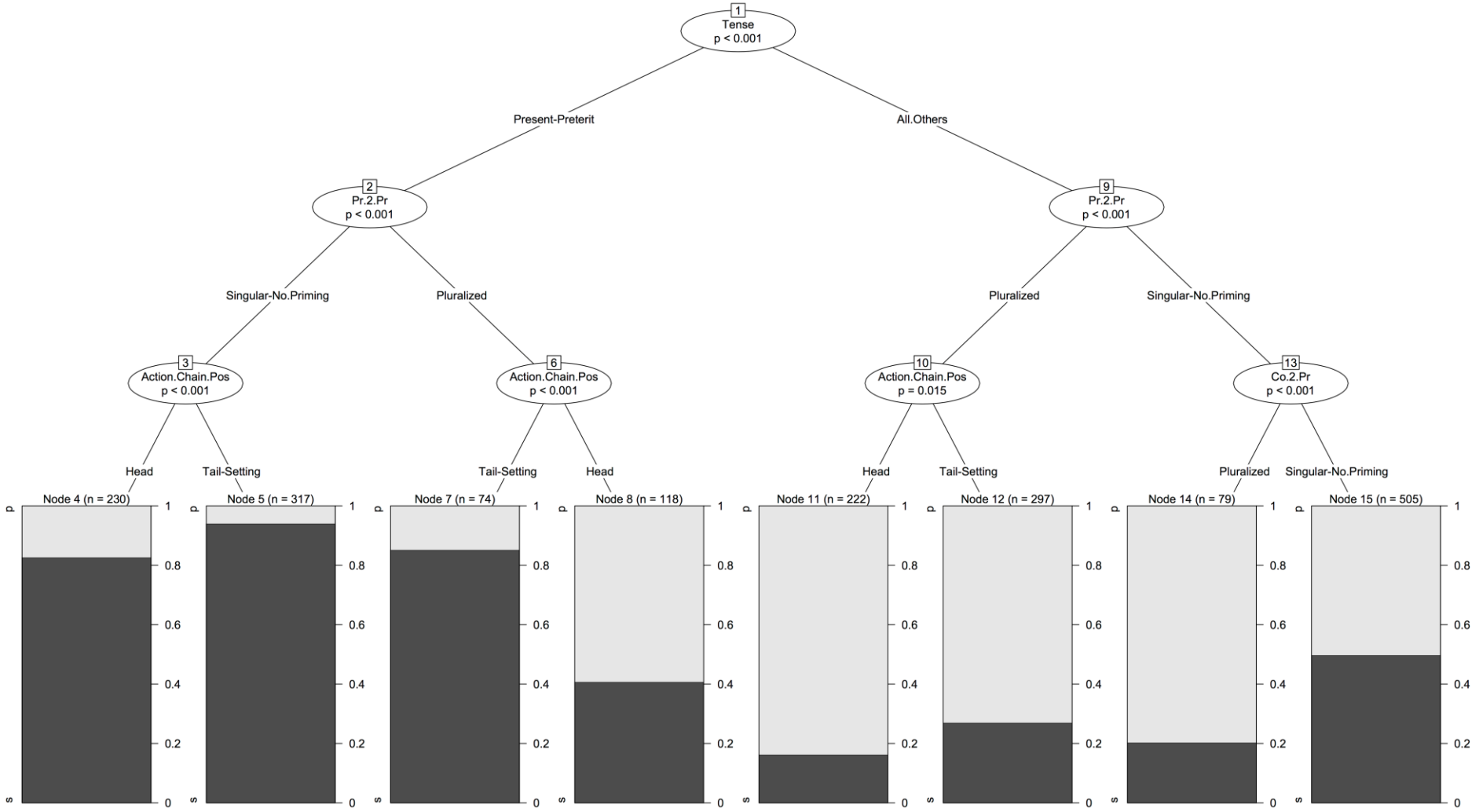
[24] *Pastel* is a Puerto Rican dish made of an outer layer of mashed plantains, filled with chicken or ground beef and vegetables. *Tembleque* is pudding made with coconut milk, sugar, and cinnamon.

Figure 6.1: Conditional inference tree model of the pluralization of presentational *haber* in Havana



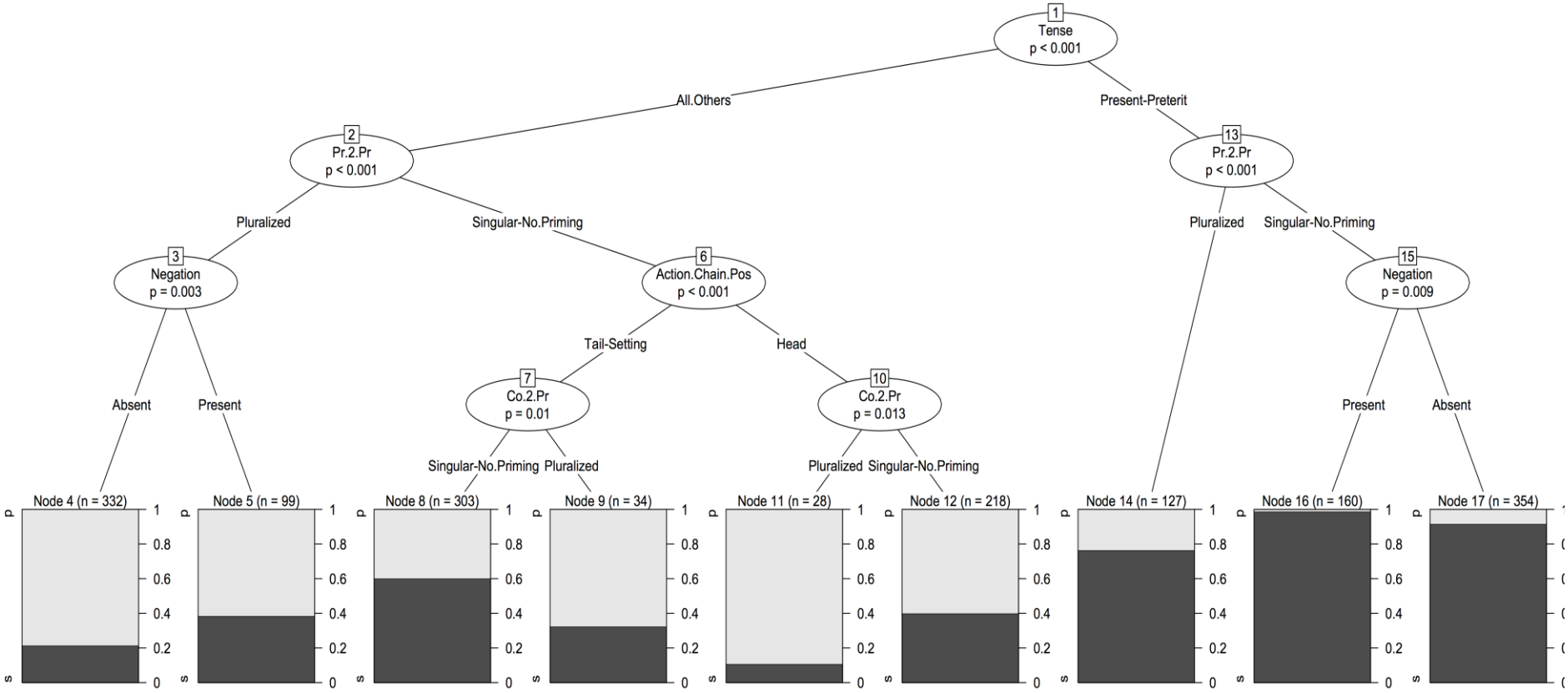
Notes: *Pr.2.Pr* means 'production-to-production priming'; *Co.2.Pr* means 'comprehension-to-production priming'; *P* means 'pluralized presentational *haber* construction'; *S* means 'singular presentational *haber* construction'

Figure 6.2: Conditional inference tree model of the pluralization of presentational *haber* in Santo Domingo



Notes: *Pr.2.Pr* means 'production-to-production priming'; *Co.2.Pr* means 'comprehension-to-production priming'; *P* means 'pluralized presentational *haber* construction'; *S* means 'singular presentational *haber* construction'

Figure 6.3: Conditional inference tree model of the pluralization of presentational *haber* in San Juan



Notes: *Pr.2.Pr* means 'production-to-production priming'; *Co.2.Pr* means 'comprehension-to-production priming'; *P* means 'pluralized presentational *haber* construction'; *S* means 'singular presentational *haber* construction'

These data suggest an antagonistic relationship in (this) language change between statistical preemption and the other two cognitive factors considered in this chapter: whereas the first encourages speakers to stick to what they have observed, the other two incite speakers to extend the pluralized presentational *haber* construction to more (and new) conceptual regions. As a result, every time the preference for unmarked coding and structural priming tip the balance in favor of the pluralized presentational *haber* construction for the encoding of a present- or preterit-tense POINTING-OUT conceptualization without aspectual or modal nuances, the use of an expression based on this construction weakens the strength of the representations of the entrenched singular instances. This, in turn, debilitates their preemptive effect, which, eventually, results in the less constrained use of <**AdvP** *hayn* **Subj**> and <**AdvP** *hubieron* **Subj**>.

The antagonistic relationship observed in this section between, on the one hand, statistical preemption and, on the other, structural priming and the preference for unmarked coding is reminiscent of the roles these cognitive factors play in language acquisition and innovation. That is, in language acquisition, statistical preemption has been shown to be the mechanism that prevents children from overgeneralizing (Goldberg, 2006a: Chap. 5, 2011), whereas structural priming has been argued to promote the extension of perceived structures to new conceptualizations of the same type (Bock & Griffin, 2000: 189; Bock et al., 2007: 455-456; Goldberg, 2009: 107; Pickering & Ferreira, 2008: 449-450). Regarding language innovation, Croft (2000: Chap. 5) argues that the tendency to maximize unmarked coding is the prime motivation for form-function reanalysis, which reforms established constructions or, put differently, overrules their preemptive effect.

Finally, the results achieved in this and the previous section also suggest that in non-experimental settings, production-to-production priming has a deeper impact than comprehension-to-production priming. In contrast, previous studies of structural priming performed under laboratory conditions found the magnitude of the priming effect to be comparable (Bock et al., 2007: 452). On a methodological note, the importance of structural priming in this variation also suggests that priming effects should not be neglected in analyses of language variation and change, even more so because psycholinguistic and variationist inquiry has shown that virtually all levels of linguistic analysis (including phonology) display priming-like phenomena (Labov, 1994: 559; Pickering & Ferreira, 2008: 429).[25]

[25] For example, Labov (1994: 559) and Poplack (1984: 213-214) observe that Puerto Ricans tend to repeat the same variant of Spanish implosive /–s/ in successive tokens. That is, "an /s/ tends to produce an /s/, and a zero tends to produce a zero" (Labov, 1994: 559).

2.5 Constraint ranking

In the previous section, it was already observed that statistical preemption and production-to-production priming emerge from both the mixed-effects regression models and the conditional inference tree models as the most important constraints on *haber* pluralization. As was explained in Chapter 4.5, the conditional permutation of predictors in a random forest model of the variation can provide more insight into this matter (Tagliamonte & Baayen, 2012: 162-164), while correcting the flaws inherent to the original range-based comparative method proposed by Tagliamonte (2002).

The results of this statistical procedure are presented in Table 6.11,[26] which shows that there are no sizeable differences between the constraint rankings of the three communities, at least if we ignore the absence/presence of negation for the moment. As hypotheses 1-3 and 9 claim that the three varieties are going through the same linguistic change, which is constrained by the same three cognitive factors, this result is entirely expected (Tagliamonte, 2002: 732-733, 2006: 246). Additionally, as was already evident from the conditional inference tree models, comprehension-to-production priming is one of the least influential constraints on the variation. As a matter of fact, only the absence/presence of negation ranks lower in San Juan.

However, in a range-based comparison, different conclusions would have been reached. Specifically, the traditional comparative method would attribute more importance to comprehension-to-production priming for Havana and San Juan, whereas this turned out to be a low-level constraint in the conditional inference tree models. For the same reason, the range-based comparison would lead us to consider the absence/presence of negation and typical action-chain position of the noun's referent as almost equally important, which is not consistent with Langacker's (1991: 312) observation that definiteness/specificity is among the least central characteristics of prototypical subjects nor with the fact that this restriction may have been loosened in Havana and Santo Domingo.

[26] For the computation of the constraint rankings, I excluded the individual speakers and lemmas, because random forest models (as conditional inference trees) tend to exaggerate the importance of grouping factors (Baayen, 2014: 366).

Table 6.11: Constraint ranking for the cognitive factors in Havana, Santo Domingo, and San Juan

Factor groups	*Havana*		*Santo Domingo*		*San Juan*	
	CPVI	*Range*	*CPVI*	*Range*	*CPVI*	*Range*
Verb tense	.0998	52	.1197	52	.1173	62
Production-to-production priming	.0441	26	.0653	31	.0265	26
Typical action-chain position of the noun's referent	.0067	12	.0255	22	.0166	16
Comprehension-to-production priming	.00007	15	.0164	21	.0046	21
Absence/presence of negation	NS	NS	NS	NS	.0010	14

Notes: *CPVI* means 'conditional permutation variable importance'. *NS* means 'not significant'

3. Social factors

Turning now to the social factors that are considered in this dissertation, Table 6.12 shows that the mixed-effects models do not consider speaker age, educational achievement, and formality to be significant constraints on *haber* pluralization. For age, this result is already implicit in the main hypothesis,[27] which, against the background of the studies reviewed in Chapter 2.2, describes the phenomenon as a slowly advancing language change from below, that is, one that might be too slow to be observed in apparent time. By the same token, hypothesis 7, which states that the rates of pluralization will not decrease as formality rises, anticipates that the alternations between the two construction schemas will not display any correlations with formality. For educational achievement, in contrast, the results do not confirm the expectations formulated in hypothesis 8. Still, the findings seem to be supported by earlier studies on Caribbean Spanish, which have shown *haber* pluralization to occur frequently in Educated Speech (DeMello, 1991; López-Morales, 1992: 147; Vaquero, 1978: 135-140). Additionally, because the effect of an extended formal education is usually described in terms of a higher sensitivity towards the formality of the encounter and a better control of formal speech styles (Labov, 1972: 138), the results for speech style and academic achievement appear to support each other.

[27] See Chapter 3.3.1.

Table 6.12: Significant social factors in the pluralization of presentational *haber* in Havana, Santo Domingo, and San Juan: Numbers, percentages, and weights for the pluralized presentational *haber* construction

Factor groups	*Havana*[a]				*Santo Domingo*[b]				*San Juan*[c]			
	N	*%*	*W(L)*	*W(S)*	*N*	*%*	*W(L)*	*W(S)*	*N*	*%*	*W(L)*	*W(S)*
Gender												
Feminine	*Not significant*				*Not significant*				375/836	44.9	.55	.55
Masculine									309/819	37.7	.45	.45
Range											10	10
Social class												
Lower	374/701	53.4	.67	.61	301/662	45.5	.47	.46				
Middle	330/700	47.1	.53	.52	248/461	53.8	.60	.60	*Not significant*			
Upper	230/692	33.2	.31	.37	311/719	43.3	.43	.44				
Range			36	24			17	16				
Speaker age: Social class (interaction group)												
20-35 years of age: lower class									125/278	45.0	.56	.56
20-35 years of age: middle class									121/273	44.3	.47	.49
20-35 years of age: upper class									132/329	40.1	.46	.47
Range											10	9
55+ years of age: middle class									104/179	58.1	.71	.70
55+ years of age: lower class									46/114	40.4	.47	.46
55+ years of age: upper class									156/482	32.4	.33	.32
Range											38	38

Notes: *W* means 'factor weight'. *(L)* means 'model with the lemmas'. *(S)* means 'model with the individual speakers'. The cognitive and social factors were included in the same regression model, but space inhibits tabulating all the results on the same page. [a] With the speakers: deviance: 2049.45; AIC: 2075.45; centered input probability: .33; with the lemmas: deviance: 2019.93; AIC: 2045.93; centered input probability: .36. [b] With the speakers: deviance: 1889.31; AIC: 1909.31; centered input probability: .46; with the lemmas: deviance: 1829.93; AIC: 1851.93; centered input probability: .45. [c] With the speakers: deviance: 1549.19; AIC: 1579.19; centered input probability: .32; with the lemmas: deviance: 1519.56; AIC: 1549.56; centered input probability: .35 (AIC = Akaike's Information Criterion)

In the three communities, the pluralization of presentational *haber* signals social class identity, although high frequencies of pluralized *haber* trigger different associations. Particularly, Table 6.12 shows that in Havana, the frequent usage of the pluralized presentational *haber* construction is associated to lower-class identity. In contrast, in Santo Domingo, it is associated to the middle class. In San Juan, in turn, social class is only significant when it is considered jointly with speaker age. That is, for older speakers, frequent *haber* pluralization signals middle class identity, whereas younger speakers associate high frequencies of pluralized *haber* with lower-class individuals. Regarding gender, Table 6.12 unveils that, in Havana and Santo Domingo, this factor does not rise above interspeaker variability. Additionally, no significant interaction can be found between gender and speaker age or any of the other social factors considered in this study. In San Juan, in contrast, the frequent usage of the pluralized presentational *haber* construction is associated to the female gender role.

These results allow evaluating whether the variation constitutes an ongoing linguistic change from below, as the main hypothesis claims.[28] In this regard, recall that the absence of a stylistic dimension is a typical feature of this type of language change (Labov, 2001: Chap. 3; Silva-Corvalán, 2001: 248-249). It should also be observed that Labov's Principles of Linguistic Change are only expected to apply uniformly to mid-way changes; some divergence from them is expected for incipient and advanced linguistic evolutions. Therefore, although the results for gender and social class do not pattern as predicted by hypothesis 5-6, this does not necessarily mean that the pluralization of presentational *haber* constitutes a stable variable or a change from above, which would imply patterns of social covariation different from the ones that emerge from Table 6.12. Specifically, if this were a stable variation or a change from above, we would expect pluralized *haber* to covary with male gender, lower social class, and informal usage events (Labov, 2001: Chap. 3). This is not the case.

Rather, thus, the findings for gender[29] and social class seem to support the idea that the three varieties have arrived at different stages of a slowly progressing, advanced language change from below, for which Labov (2001: 308-309) observes that gender differences tend to become smaller or disappear altogether and that a linear alignment with social class may develop. In this sense, the results of this study seem to corroborate those of earlier investigations on Latin American Spanish (D'Aquino-Ruiz, 2008; Díaz-Campos, 2003; Fontanella de Weinberg, 1992b), while at the same time supporting the main hypothesis and hypothesis 9. Let us review now the most important findings of this chapter.

---

[28] See Chapter 3.3.1.

[29] That is, no gender stratification in Havana and Santo Domingo and covariation with female gender in San Juan.

4. Summary

This chapter has presented a series of quantitative analyses designed to test the hypotheses introduced in Chapter 3.3. First, a more in-depth analysis of the Puerto Rican dataset was performed, which was intended to single out the relevant constraint pertaining to the NP. The results unveiled that the typical action-chain position of the noun's referent is the feature to which *haber* pluralization is sensitive. This parameter was subsequently applied to the three corpora, which showed that typical action-chain heads favor *haber* pluralization. In addition, it was argued that the presence of negation contributes to portraying the NP as a less prototypical subject. Indeed, in San Juan, speakers are less likely to select the pluralized presentational *haber* construction when negation is present. For Havana and Santo Domingo, in turn, the absence/presence of negation did not turn out to be a significant constraint on the variation. As these varieties also display slightly higher overall rates of pluralization, these findings may suggest that the pluralized presentational *haber* construction has invaded the nonspecific indefinite conceptual region in Havana and Santo Domingo. Still, the data for this and the previous factor group support the first hypothesis: speakers tend to encode more prototypical subjects as subjects with the pluralized presentational *haber* construction. Regarding the influence of the verb tense, this chapter has shown that the tendency to pluralize *haber* less often in synthetic present- and preterit-tense expressions supports hypotheses 2a-c, as these forms were used predominantly in (singular) presentational *haber* clauses before the change took off. Additionally, it was shown that speakers are more likely to pluralize *haber* when they have just processed or used a pluralized presentational *haber* construction, regardless of variations in tense, aspect, or mood. These results support the third hypothesis while also suggesting that *haber* occurs in two conventionalized construction frames. Exploring the interaction between the three cognitive constraints, in turn, has revealed that structural priming and the preference for unmarked coding incite speakers to extend the use of the pluralized presentational *haber* construction to the present and the preterit tense, whereas statistical preemption works against this. Finally, contrasting the constraint rankings of the three varieties, we have seen that the three cognitive factors essentially have identical effects, with one exception: the absence/presence of negation.

Turning now to the social factors considered in this investigation, the mixed-effects regression models show a linear social class alignment for Havana and the youngest age group of San Juan. In contrast, for the oldest age group of this community and for Santo Domingo, Rbrul indicates that *haber* pluralization is associated to the middle class. Concerning gender, the data display no gender differences for Havana and Santo Domingo, whereas in San Juan, high rates of *haber* pluralization are interpreted as signaling female gender. These data support the conclusion that *haber* pluralization

constitutes an advanced language change from below. In the next chapter, these results will be used to argue for a novel account of the emergence of the pluralization of presentational *haber*.

7

A Cognitive Construction Grammar approach to the emergence of (pluralized) presentational haber

Chapter 6 has revealed that in Caribbean Spanish, the pluralization of presentational *haber* constitutes an advanced linguistic change from below that is constrained by three general cognitive factors. In section 2, this information will be used to argue in favor of a novel account of the emergence of *haber* pluralization. However, first, it might be helpful to explore how the emergence of the singular presentational *haber* construction could be modeled in Cognitive Construction Grammar. This will be the focus of section 1.

1. A Cognitive Construction Grammar approach to the emergence of presentational *haber*

This section presents a Cognitive Construction Grammar approach to the emergence of presentational *haber*. Specifically, the first section introduces the Classical Latin presentational construction with *esse* 'to be' and its Late Latin variant with *habere* 'to have'. In the second section, possessive *habere* and *have*-type possessives in general will be discussed. In section 3, this will lead to the insight that a bleached variant of possessive *habere* could easily be reanalyzed as a presentational expression.

1.1 Classical Latin presentational esse *and Late Latin presentational* habere

In Classical Latin, presentational expressions were most commonly formed with presentational *esse* 'to be', as in example (1) (Hernández-Díaz, 2006: 1125; Herrero-Ruiz de Lozaiga, 2008: 341; Luque-Moreno, 1978: 140; Moreno-Bernal, 1978: 283). This presentational construction schema was also instantiated by verbs that expressed a precondition to existence, such as, for instance, *exsistere* 'to emerge, to come forward, to arise' in example (2). The second example also shows that, like present-

day presentational *haber*,[1] the Classical Latin presentational construction included a profiled adverbial phrase (the ablative *hoc loco* 'at this point'), which could remain implicit when it referred to the base space of the usage event or a previously evoked mental space, as in example (1). Since this construction was used to introduce new referents into discourse, it was probably also subject to the same information-status constraints as present-day presentational *haber* (see Figure 7.1).

(1) *Erant* complures honesti adulescentes, senatorum filii et ordinis eques tris; *erant* legationes civitatum, *erant* legati Caesaris (Caesar, *De bello civili,* 1.51, 1st century B.C.).
'*There were* a number of honorable young men, sons of senators and the equestrian order; *there were* deputations from the cities, *there were* lieutenants of Caesar's.'

(2) *Exsistit* autem hoc loco quaedam quaestio subdifficilis (Cicero, *De Amicitia*, 67, 1st century B.C.).
'*There arises*, however, at this point a somewhat difficult question.'

Figure 7.1: The Classical Latin presentational construction

Sem	POINTING-OUT	<	**location**	**zero**	>
R: instance,	|R		|	|	
precondition	PRED	<			>
	↓		↓	↓	
Syn	V		adverbial phrase	subject	
Prag			-	hearer-new	

In Late Latin, however, a novel presentational schema with the, originally, possessive verb *habere* 'to have' replaces this construction (García-Yebra, 1983; Herrero-Ruiz de Lozaiga, 2008: 341; Luque-Moreno, 1978: 135-136; Väänänen, 1967: 136-137). In section 1.3, a reconstruction of this evolution will be attempted, but let us first review the characteristics of Classical Latin possessive *habere* and *have*-type possessives in general.

1.2 Classical Latin habere *and* have-*type possessives*

In Classical Latin, two types of clausal possessive constructions can be identified: an older *be*-type possessive with *esse* (see example 3) and a *have*-type construction, typically instantiated by *habere* (see example 4) (Gónzalez-Calvo, 2002: 644; Lyons, 1967: 391-392), which progressively replaces the first one (Heine, 1997: 109; Lyons, 1967: 392). The possessive with *esse* encodes the possessed goods as subjects and the possessors as goals, inflected in the dative case (Clark, 1978: 115; González-Calvo,

[1] See Chapter 5.3-Chapter 5.4.

2002: 644; Heine, 1997: 61; Lyons, 1967: 392). In the possessive construction with *habere*, in turn, the possessed goods function as objects, while the possessors are assigned the subject role.

(3) *Est mihi* nonum superantis annum plenus Albani cadus (Flaccus, *Carmina*, 4.11.1, 1st century B.C.).
'*I have* a cask full of Albanian wine older than nine years old.'

(4) Si *habeo* familiam, a familia mea fateor te esse deiectum (Cicero, *Pro A. Caecina*, 55.13, 1st century B.C.).
'If *I have* a household of slaves, I admit that you have been sent off by my household of slaves.'

According to Langacker (1991: 170-180, 1995: 58-60, 1999: 176, 2009: Chap. 4), the schematic meaning of both types of possessive constructions can be characterized in terms of the 'reference-point ability', in other words, our ability to establish mental contact with an entity through a reference point (Langacker, 2009: 82). More precisely, Langacker (1991) argues that all possessive expressions (including nominal possessives) minimally indicate that

> the conceptualizer traces a mental path through the reference point to the target; the reference point constitutes the possessor, and the target, the entity possessed (Langacker 1991: 171).

Possessive *esse* encodes this ICM quite transparently, in the sense that this construction combines a stative verb with a reference point in the dative case, often used for locations across languages (Clark, 1978: 115). For *habere*-possessives, however, encoding a stative relationship with a transitive construction constitutes a deviation from prototypical transitive clauses, which refer to energetic events rather than situations (García-Yebra, 1983: 60; Hernández Díaz, 2006: 1060; Langacker, 1991: Chap. 7, 2009: 93; Luque-Moreno, 1978: 132-133; Lyons, 1967: 392). The same is true for English clausal possessives with *to have*, as is evident from example (5).[2]

(5) I *have* a big collection of Lacoste shirts that are now vintage (Davies, 2008-, Press).

This suggests that *have*-type possessives instantiate a specialized possessive argument-structure construction, represented in Figure 7.2. Since this construction schema encodes the possessive ICM, its subject can arguably be considered agentive. Rather, because the subject merely specifies a reference point, it fulfills a location role (Clark, 1978: 116-117). The object, in turn, refers to a target entity that is simply present in the picture presented through the construction, for which it is probably most

[2] This also applies to other stative, transitive predicates such as e.g., *to love, to hate,* etc. See Langacker (1991: Chap. 8) for discussion.

adequately conceptualized as a 'zero' argument,[3] which, schematically, functions as trajector. Since this possessive construction is used to point out the existence of the object/target entity and its relationship to the reference point, both the target entity and its relation to the reference point constitute the new information that is communicated to the hearer.[4] In contrast, the subject/reference point may communicate both new (see example 6) and known information (see example 7).

(6) Just because a seller *has* jalapeños in a tray that sells for $3 doesn't mean you have to buy $3 worth (Davies, 2008-, Press).

(7) The Chinese *had* gunpowder, but failed to arm their troops with guns (Davies, 2008-, Magazine).

Figure 7.2: The *have*-type possessive construction

Sem	POSSESSIVE	<	**location**	**zero**	>
R: instance,	|R		|	|	
precondition	PRED	<			>
	↓		↓	↓	
Syn	V		subject	object	
Prag			-	hearer-new	

As noted earlier, the reference-point ICM is the semantic common denominator that unites all instances of this construction schema, but this does not keep specific verbs from adding a sense of agentive control by the reference point over the target entity (Langacker, 1991: 171, 1995: 64, 1999: 182). From Langacker's (1999: 182) perspective, these more agentive possessive expressions do not contradict the subject's schematic role as a reference point, because controlling an object also implies that the controller is aware of its location and that the controlled entity can be accessed through her/him. In addition, because the 'zero' argument role is subsumed in all others (Langacker, 1991: 288), it can be instantiated by more patient-like participant roles, such as, for example, those listed in (8) for the verb *to hold*.

(8) *To hold* <**holder**, **holded**>.

In Figure 7.2, the instance and precondition relationships capture this possibility. For example, the instance relationship accounts for the possessive uses of verbs such as *to hold* (see example 9), which denote physical control of the reference point over the target entity.

(9) She *holds* a radio in her left hand, against her thigh (Davies, 2008-, Fiction).

[3] See Langacker (1991: 288).

[4] As a consequence, the target entity is often indefinite (Heine, 1997: 31).

The precondition relationship, in turn, captures the fact that verbs such as *to grasp* in example (10) can be used to express possession, because they refer to an action resulting in the situation described by the possessive ICM.

(10) In each hand he *grasps* a golden staff with a floriated finial (Davies, 2008-, Academic).

Additionally, the idea of physical control of the reference point over the target entity may also emerge from context (Heine, 1997: 87). This is most evident for bleached-out verbs such as English *to have,* which can express a wide range of possessive relationships (Langacker, 1995: 64). For instance, in example (11), without the ensuing context, the *have* clause merely denotes "ownership implying the possibility of physical access whenever desired" (Langacker, 1995: 64). In contrast, with the right-hand context, the expression indicates that the reference point is physically controlling the target entity, a meaning that no longer emerges out of context with this verb (Bybee & Pagliuca, 1985: 71).

(11) "My brother *has* a knife, and he's trying to stab my mother with it," she told the 911 operator again, seconds later, in the same flat voice as before (Davies, 2008-, Press).

1.3 From Classical Latin possessive habere *to Late Latin presentational* habere

Like English *to have* (Bybee & Pagliuca, 1985: 71; Heine, 1997: 47-50; Langacker, 1991: 172), Latin *habere* evolved from a prototypical transitive verb, meaning 'to hold' into a more stative verb expressing the possessive ICM (Heine, 1997: 109-110). As the most frequently used instance of this construction schema, *habere* was probably stored mentally as one of its special cases, which, originally, included at least a vestige of active control by the reference point over the target entity. Indeed, examples such as (12) show that, in the first century B.C., it is still possible to document *habere* with this meaning.

(12) Quod mustum, conditur in dolium ut *habeamus* vinum (Varro, *Res Rusticae*, 1.65.1.1, 1st century B.C.).
'The must, which is put into a jar such that we *have/obtain* wine.'

This suggests that, in Classical Latin, *habere* featured participant roles that instantiated the argument roles of the construction in a more specific way, namely 'possessor' and 'possessee' (see Figure 7.3). These roles, in turn, disfavored the use of the verb with inanimate subjects incapable of physically controlling a possession.

Figure 7.3: The Classical Latin entrenched possessive *habere* instance

Sem	POSSESSIVE	<	**location**	**zero**	>
R: instance	|R		|	|	
	habere	<	**possessor**	**possessee**	>
	↓		↓	↓	
Syn	V		subject	object	
Prag			-	hearer-new	

However, it appears that by the Late Latin period, this entrenched instance of *habere* had been bleached (Luque-Moreno, 1978: 139-140; Stengaard, 2013: 212, 220) to the same extent as present-day English *to have*. This is supported by the fact that in Late Latin and Old Ibero-Romance, agentive uses of *habere* are scarce (Stengaard, 2013: 220) and that *tenere* ('to hold', 'to subject', 'to maintain') becomes more widely used to express the physical control of animate possessors over possessions (Garachana-Camarero, 1997: 222; Heine, 1997: 109-110; Hernández-Díaz, 2006: 1108-1109). The fact that, in Late Latin, *habere* came to be used in novel aspectual, future-tense, and modal auxiliary constructions (Garachana-Camarero, 1997: 213; Heine, 1997: Chap. 4; Luque-Moreno, 1978: 132-133, 145; Norberg, 1968: 24; Penny, 2006: 193-194, 237; Väänänen, 1967: 140-141) also implies that its meaning had already become more schematic before that time (Bybee, 2003a: 152; Bybee & Pagliuca, 1985: 71).

One of the consequences of bleaching is that the more specific participant roles of verbs come to coincide completely with the abstract argument roles specified by the construction schema they are entrenched in. This typically leads to a situation in which "expressions for human concepts come to be used also for concepts that are inanimate" (Heine, 1997: 87). For *habere,* the bleaching of its semantics implied that the verb dropped the requirement of a reference point capable of physically controlling a target entity. As a result, the Late Latin entrenched *habere* instance (see Figure 7.4) could accept all types of subjects, provided they could be construed as references points. This is demonstrated in example (13), where a location is used as subject with possessive *habere*.

(13) Aecclesia *habet* de se gratiam grandem (Aetheria, *Peregrinatio Aetheriae,* 3.3, 4th century. From Stengaard, 2013: 216).
'The church *has* great beauty from its own.'

Figure 7.4: The Late Latin entrenched possessive *habere* instance

Sem	POSSESSIVE	<	**location**	**zero**	>
R: instance	|R		|	|	
	habere	<	**setting**	**participant**	>
	↓		↓	↓	
Syn	V		subject	object	
Prag			-	hearer-new	

As Langacker (1995: 74) observes, possessive expressions with location subjects "verge on being existential". Indeed, examples such as (13) do little more than setting up a mental space (*the church*) and pointing out one of its qualities, for which this example is almost synonymous with its presentational paraphrase provided in example (14).[5] For this reason, form-function reanalysis of possessives with location subjects as presentationals is likely to occur (Langacker, 2009: 107).[6]

(14) *There is* great beauty in the church itself.

In Late Latin, the reanalysis of expressions such as (13) appears to have been favored by three factors. To start with, *habere* had always had some intransitive uses in the spoken language, signifying 'to live in' (see example 15) or attributive 'to be' (see example 16) in fixed expressions (García-Yebra, 1983: 60; Hernández-Díaz, 2006: 1125; Luque-Moreno, 1978: 143-145; Weber, 1985: 285-286). The erosion of the case system would have favored the process as well, as it further obscured the syntactic status of the reference point (Norberg, 1968: 23-24; Väänänen, 1967: 117-124). The increased contact with Greek through literal translations of biblical texts has also been suggested as a contributing factor (González-Calvo, 2002: 642; Hernández-Díaz, 2006: 1125; Norberg, 1968: 15-17).

(15) Qui Syracusis *habet* (Plautus, *Menaechmi*, 68-69, 1st century B.C. From Hernández-Díaz, 2006: 1125).
'Who *lives* in Syracuse.'

(16) Bene *habet* (Cicero, *Pro Murena*, 14, 1st century B.C. From Hernández-Díaz, 2006: 1125).
'*It's* good.'

Technically, the reanalysis of bleached possessive *habere* came down to speakers encoding the typologically marked subject in a less marked fashion, that is, as an adverbial phrase. As the object of this clause type did not deviate significantly from its

[5] This illustrates the often-observed conceptual similarities between presentational and possessive clauses (Clark, 1978; Freeze, 1992; Lyons, 1967).
[6] See Chapter 3.1.8.

prototypical value,[7] form-function reanalysis did not affect the construction further. The result was the subjectless, impersonal presentational *habere* construction illustrated in example (17) and represented in Figure 7.5. From the beginning, this novel construction could be used either to point out the existence of some entity in a location (see example 18) or to point out the distance in time or space between two entities (see example 19) (Luque-Moreno, 1978: 136).

(17) In arca Noe *habuit* homines (Saint Jerome, *Epistolae*, 123.9, 4th-5th century. From Herrero-Ruiz de Lozaiga, 2008: 341).
'In Noah's arc *there were*$_{Sing}$ men.'

(18) *Habet* in Bibliotheca Ulpia in armario sexto librum elephantinum (Flavius Vopiscus, *Historia Augusta: Vita Taciti Imperatoris*, 8.1, 4th century. From Väänänen, 1967: 137).
'*There is* in the Ulpian Library, in the sixth case, an ivory book.'

(19) *Habebat* autem de eo loco ad montem Dei forsitan quattuor milia (Aetheria, *Peregrinatio Aetheriae*, 1.2, 4th century. From Luque-Moreno, 1978: 136).
'From that place to God's Mountain *there were*$_{Sing}$ some four miles.'

Figure 7.5: The Late Latin singular presentational *habere* construction

Sem	POINTING-OUT	<	**location**	**zero**	>
R: instance	|R		|	|	
	habere	<	**setting**	**participant**	>
	↓		↓	↓	
Syn	V		adverbial phrase	object	
Prag			-	hearer-new	

Once conventionalized, this construction did not replace the bleached possessive *habere* schema nor the original presentational *esse* construction right away. Rather, these three constructions would continue to coexist until the fifteenth/sixteenth century (Garachana-Camarero, 1997; Hernández-Díaz, 2006; Herrero-Ruiz de Lozaiga, 2008; Moreno-Bernal, 1978: 283-284), the timeframe that will concern us in the next section.

[7] In the sense that it is typically inanimate, indefinite, and hearer-new (see Langacker, 1991: Chap. 7).

2. Some speculations on the emergence of pluralized presentational *haber*

Ultimately, the singular presentational *haber* construction evolved into the pluralized presentational *haber* construction, implying an additional reanalysis, this time of the *presentatum* as subject. However, the mechanism and the timeframe of this evolution remain subject to debate. In gross strokes, three main lines of argumentation can be identified. The most recent perspective is that of Hernández-Díaz (2006: 1048), who favors the view that the pluralization of *haber* corresponds to a relatively recent pan-Hispanic innovation. Fontanella de Weinberg (1987, 1992b), for her part, claims that *haber* pluralization constitutes a gradually progressing syntactic change that appears around the eighteenth century in American Spanish. Still, the majority view appears to be that the phenomenon was already present in Late Latin or Old Spanish (Bentivoglio & Sedano, 1989; García-Yebra, 1983; Luque-Moreno, 1978; Moreno-Bernal, 1978).

Whichever may be the line of thought that one wishes to adhere to, some facts about the diachronic development of *haber* pluralization should not be ignored. To begin with, as the proponents of the majority view have argued repeatedly, already in Late Latin (see example 20), Old French (see example 21), and Old Spanish (see example 22), occasional examples of nominative case endings and/or verb agreement with presentational *habere* or its Romance descendants can be found (García-Yebra, 1983: 71; Luque-Moreno, 1978: 146; Moreno-Bernal, 1978: 291; Stengaard, 2013: 218-219).

(20) Piscina Siloe a lacu, ubi missus est Hieremias propheta, *habet* passus numero C (Theodosius*, De situ Terrae sanctae*, 6th century. From Luque-Moreno, 1978: 146).
'From the Fountain of Siloam to the pool, where Jeremiah the prophet was put, *there is* a number of one hundred passes$_{Nom}$.'

(21) Uns almaçurs *i ad* de Moriane, *n'ad* plus telun en la iere d'Espagne (Anonymous, *La Chanson de Rolande*, 11th century. From García-Yebra, 1983: 71).
'*There is* an almaçour$_{Nom}$ of Moriane, *there isn't* a more villain one$_{Acc}$ in the Spanish land.'[8]

(22) E *avién y* grand abondo de buenos arboledas (Gónzalo de Berceo, *Milagros de Nuestra Señora*, 13th century. From Stengaard, 2013: 219).
'And *there were*$_{Plur}$ great abundance of good woodlands.'

[8] 'Almaçour' was the title for a provincial chief in Islamic Iberia.

Table 7.1: Early examples of pluralized presentational *haber*

Century	*Example*
13th	E *avién* allí muchos engeños e muchas armas (from Moreno-Bernal, 1978: 291). 'And there, *there were*$_{Plur}$ a lot of scams and a lot of weapons.'
15th	Algunos *ovieron* que o con mala voluntad, o no discretamente, quisieron disfamar al rey de Navarra e al Infante don Enrrique (from González-Calvo, 2002: 652). '*There were*$_{Plur}$ a few who, or out of bad will, or not discretely, wanted to defame the king of Navarra and the Crown Prince Henry.'
16th	Acá *an abido* ciertas reboluciones (from Frago-Gracia, 1999: 112). 'Here, *there have been*$_{Plur}$ certain revolutions.'
16th	En esta flota que vino de España pensé que *hubieran* algunas cartas (from Fontanella de Weinberg, 1992a: 70). 'In this fleet that came from Spain, I thought that *there would be*$_{Plur}$ some letters.'
17th	En ella *hubieron* cosas dignas de memoria… (from Kany, 1945/1951: 256). 'In that one, *there were*$_{Plur}$ things worthy of remembrance...'
17th	En la boda *hubieron* danzas (from González-Calvo, 2002: 653). 'At the wedding, *there were*$_{Plur}$ dances.'
18th	Otras mujeres *habían* honestas (from Fontanella de Weinberg, 1992a: 70). 'Other women *there were*$_{Plur}$ honest ones.'

As a matter of fact, throughout the history of Spanish, cases of pluralized presentational *haber* have been documented (Gómez-Torrego, 1994: 31; Kany, 1945/1951: 256; Moreno-Bernal, 1978: 290-291; Quintanilla-Aguilar, 2009: 53; Stengaard, 2013: 218-219; Suñer, 1982: 101), as is illustrated in Table 7.1. Yet, at the same time, the results of multiple diachronic studies also suggest that pluralized presentational *haber* is scarce before the eighteenth century (Álvarez-Nazarío, 1991: Chap. 4; Fontanella de Weinberg, 1987: Chap. 1-2, 1992a: 70, 1992b; González-Calvo, 2002: 655; Gutiérrez-Grova, 2007; Hernández-Díaz, 2006: 1146-1152).

Second, as we have seen in Chapter 2.1, the pluralization of presentational *haber* can occasionally be found in nearly every dialect of Spanish, but it is only in the coastal areas of the Spanish Mediterranean, the Canary Islands, and the Americas that robust variability has been documented. In addition, any hypothesis on the emergence of *haber* pluralization should also be able to account for the fact that Catalan *haver-hi* displays similar patterns of variable agreement in certain dialects (Blas-Arroyo, 1995-1996, 1999: Chap. 2; Ramos-Alfajarín, 2001; Rigau, 1993).

Taking into consideration the three points raised so far, it is rather doubtful that the examples in Table 7.1 and the occasional pluralized tokens that can be found in non-pluralizing areas point to a frequently occurring phenomenon, indicative of a linguistic

change (González-Calvo, 2002: 651). Rather, these cases can be taken to suggest that, throughout the history of Spanish, speakers have always experienced the singular presentational *haber* construction as an exceptional pattern (Fontanella de Weinberg, 1992b: 40; Hernández-Díaz, 2006: 1127-1129; Kany, 1945/1951: 256; Real Academia Española & Asociación de Academias de la Lengua Española, 2009: §41.6c; Suñer, 1982: 106).

Indeed, the syntax of this construction finds little motivation in the overall system of Spanish nor in its meaning. Concerning the first factor, besides presentational *haber*, Spanish has only one other one-argument subjectless impersonal construction without overt impersonality marker, namely, the presentational *hacer* construction shown in example (23) (Fontanella de Weinberg, 1992b: 40). As the dominant trend in Spanish, thus, clearly consists in using overt markings to indicate subjectless impersonality, these two construction schemas, both confined to just one verb and a specific type of NP argument, have always been prone to form-function reanalysis (Fontanella de Weinberg, 1992b: 40). As a matter of fact, presentational *hacer* also displays agreement variation in Canarian and Latin American Spanish (Álvarez-Nazarío, 1991: 490, 709; Bentivoglio & Sedano, 1996: 124; Catalán, 1989: 226; Fontanella de Weinberg, 1987: 108, 1992a: 70, 1992b: 40).[9]

(23) Fue en el año, *hace* como dos años que ya no hay ciclón (LH07M11).
'It was in the year, *it's been*$_{Sing}$ like two years that there isn't a hurricane.'

Regarding the second factor, Langacker (1991: Chap. 7) argues that the semantic feature that unites all subjects is that of being the primary figure at a clausal level of organization. As the NP of presentational *haber* fulfills just this role in the POINTING-OUT ICM,[10] the singular presentational *haber* construction represents a tension between the NP's cognitive salience and its syntactic encoding as an object (Kany, 1945/1951: 256; Suñer, 1982: 106). In this light, sporadic form-function reanalysis can also be expected to occur (Croft, 2000: Chap. 5).

While this shows that singular presentational *haber* has always been prone to form-function reanalysis, it does not explain why *haber* pluralization is only predominant in the Spanish of Eastern Spain, the Canary Islands, and Latin America.[11] In this regard, it should be observed that these areas all have welcomed large-scale population movements in the past, which are known to trigger linguistic changes (Labov, 2001: Chap. 9; 2010: Chap. 5). Therefore, it is at least partially predictable that, in these areas, a poorly motivated pattern such as the singular presentational *haber* construction

[9] However, since the occurrence rate of presentational *hacer* is drastically lower than that of presentational *haber*, it is virtually impossible to study the pluralization of *hacer* in a corpus of sociolinguistic interviews.
[10] See Chapter 5.1.
[11] See Chapter 2.1.

would begin to shift, leading to an ongoing linguistic change from below. For the Canary Islands and Latin America, the most likely candidate for such a triggering event appears to be the colonization process (fifteenth/sixteenth centuries), which involved intense contacts between the dialects of settlers from all over Castile,[12] the languages of the indigenous populations, and those of transplanted African slaves (Álvarez-Nazarío, 1991: Chap. 2, 6, 7; Boyd-Bowman, 1976; De Granda, 1994: Chap. 1; Lipski, 2009; Lope-Blanch, 1989: 12; López-Morales, 1998: Chap. 1, 2, 5; Lüdtke, 1994: 43-44; Medina-López, 1999: Chap. 2; Ortiz-López, 2000: 364-365). The result of this intense dialect and language contact was the emergence, by the end of the sixteenth century, of new Canarian and Latin American contact varieties, of which the present-day Canarian and Latin American dialects are the direct descendants (De Granda, 1994: Chap. 1; Fontanella de Weinberg, 1992a: 44-47).

The ethnolinguistic profiles of the immigrants to the Spanish colonies suggest that in the emergence of this new variety, two simultaneous processes were at play: koineization and language contact through adult language learning. For the first process, Siegel (1985) provides the following definition:

> [k]oineization is the process, which leads to mixing of linguistic subsystems, that is, of language varieties, which either are mutually intelligible or share the same genetically related superposed language. It occurs in the context of increased interaction or integration among speakers of these varieties (Siegel, 1985: 375-376).

As is evident from this description, koineization is a consequence of long-term mutual accommodation, in other words, of speakers of different dialects of the same language adapting their speech to achieve better communication with one another (Kerswill, 2002: 680; Siegel, 1985: 367; Trudgill, 1986: Chap. 1). A 'koine', then, "is the stabilized composite variety, which results from this process" (Siegel, 1985: 375-376).[13]

According to Trudgill (1986: Chap. 3), koineization involves two simultaneous processes ('leveling' and 'simplification'), which typically take two to three generations to complete (Kerswill, 2002: 670; Trudgill, 1986: 98). The first process, leveling, indicates that the koine only retains the dialectal variant used by the majority of the speakers of the community (Kerswill, 2002: 671-675; Kerswill & Williams, 2000: 85; Trudgill, 1986: Chap. 3). Simplification, in turn, indicates that in the process of leveling, the koine will not select the variant of the majority dialect when the alternative provided by another source dialect implies a reduction of the total number

[12] By 1492, Spain was a personal union between the Catholic Kings Isabel of Castile and Fernando of Aragón, who, although married to each other, were sovereigns in their separate kingdoms (Cruz-Arroyo, 2000: 257-258). The colonization effort was funded by Castile and citizens of Aragón did not emigrate frequently (Boyd-Bowman, 1976).

[13] Of course, this implies that speakers actively try to establish social networks across dialect borders (Kerswill, 2002: 673).

of phonological distinctions or morphosyntactic categories within the system (Kerswill, 2002: 671-674; Kerswill & Williams, 2000: 67, 85; Trudgill, 1983: Chap. 5-6, 1986: 103). As a result of systemic pressures,[14] simplification may also produce novel, interdialectal compromise forms (Trudgill, 1986: 110). Either way, simplification leads to "an increase in morphological and lexical transparency" (Trudgill, 1986: 103).[15]

As noted, koineization results in a virtually homogeneous koine variety, with which the newly established community identifies (Kerswill, 2002: 689-695; Trudgill, 1986: Chap. 3). However, this does not mean that variation is necessarily absent once the koine has stabilized. Rather, as Siegel (1985: 375) observes, "'rekoineization' can take place if there is continued contact with the original closely related varieties, or additional contact with different ones". Additionally, even without rekoineization, some residual variation may continue to exist, in which case the variants are refunctionalized as socially and/or stylistically distinct alternatives (Trudgill, 1986: 126).

For Latin America, De Granda (1994: 36-37) and Fontanella de Weinberg (1992a: 44-47) argue that throughout the colonial period, koineization and rekoineization have occurred. Specifically, these authors show that of the different variants that were brought to Latin America, the koine consistently selected the one that implied a reduction of the total number of distinctions in the language system (De Granda, 1994: Chap. 1-2). Since the majority of the immigrants were Andalusian (Boyd-Bowman, 1976), the variety had a strong Andalusian flavor, but it adopted northern features provided these implied a reduction of the number of phonological or morphosyntactic categories (Catalán, 1989: 142; De Granda, 1994: 34-37, 70). A similar process must have taken place on the Canary Islands (Catalán, 1989: 124; Medina-López, 1999: 55).

While this accounts for the features of Canarian and Latin American Spanish, their resemblance to one another, and their similarity to the Andalusian varieties of Peninsular Spanish, the colonial history of the Canary Islands and Latin America suggests that language contacts through adult language learning may also have shaped the koine. In this regard, it should be observed that before 1540-1560, the emigration of women and children to the Americas was virtually inexistent (Boyd-Bowman, 1976: 582). Even in later years, emigrating to the colonies remained predominantly a young

[14] For example, the preference for unmarked coding.

[15] However, extreme simplification only appears to take place when the bulk of the dialect contact occurs, between adults, as was the case in the early Spanish colonies (Boyd-Bowman, 1976). When important numbers of children are added to the equation, as in the British New Town of Milton Keynes, the resulting koine variety displays little to no simplification (Kerswill & Williams, 2000: 89). This is another example of the long-standing observation in sociolinguistics that "a native-like command of a linguistic pattern different from that first learned is possible only for children who move into the new community before the age of nine or ten" (Labov, 2010: 8).

man's game (López-Morales, 1998: 45-46). As a consequence, Spanish colonists mixed with indigenous women (López-Morales, 1998: 28), producing a Spanish-dominant bilingual society (López-Morales, 1998: Chap. 2). Additionally, from 1518 onward, the Spanish began introducing African slaves, who were mainly employed as domestic servants in urban settings or on small farms (Álvarez-Nazarío, 1991: 357; Lipski, 2009: 48-50, Chap. 4; Moya-Pons, 2008: Chap. 2; Picó, 2000: 62). Since slaves lived and worked in close contact with the Spanish population,

> blacks were in constant contact with local varieties of Spanish … , although American-born blacks may have retained certain ethnolinguistic markers as a consequence of their inevitably marginalized status (Lipski, 2009: 49). [16]

On the Canary Islands, an additional language contact existed with Portuguese (Corbella-Díaz, 1996: 115-122; Lüdtke, 1994: 44; Medina-López, 1999: Chap. 2).

This suggests that in the early Canarian and Latin American settlements, language contacts were probably as prevalent as dialect contacts. In this light, Africanized, Native-American, Native-Canarian, and Portuguese L2 varieties of Spanish can be expected to have contributed to at least some degree to the mix of varieties from which the Canarian and Latin American koines arose (Álvarez-Nazarío, 1991: 356-357). Indeed, in the Canarian and Latin American lexicons, multiple contributions of African and indigenous languages can be found (Álvarez-Nazarío, 1991: Chap. 6-7; Corbella-Díaz, 1996; López-Morales, 1998: 30-41, 96-103; Lüdtke, 1994: 44; Medina-López, 1999: 54-55). In the Canarian lexicon, Portuguese loans have also been identified (Corbella-Díaz, 1996: 115-122; Lüdtke, 1994: 45; Medina-López, 1999: 54).

---

[16] Spanish-based creole languages never developed (see e.g., Fontanella de Weinberg, 1992a: 241-249; Lipski, 1996, 2009; Megenney, 1999 and Ortiz-López, 2000: 366-372 for discussion). This was mainly due to the fact that, as already noted, African-born forced laborers (called '*bozales'*) lived in close contact with European settlers. In addition, due to the relatively low demand for forced labor, the import of new slaves from Africa was rather limited, because the natural renewal of the population was sufficient to fulfill the needs (Lipski, 2009: 10, Chap. 4). Also, the original slave population had not been imported directly from Africa, but rather from Spain, where they had already learned (some) Spanish (Álvarez-Nazarío, 1991: 361; López-Morales, 1998: Chap. 5; Moya-Pons, 2008: 184; Valdés-Bernal, 1994: 7, 2007b: 41). These three factors ensured that slaves had enough contact with Spanish to acquire the language, although with varying proficiency (Álvarez-Nazarío, 1991: 363; Lipski, 2009: Chap. 4). However, when the plantation economy finally took off in Cuba and Puerto Rico, after the Haitian Revolution of 1790 (Moya-Pons, 2008: Chap. 15), the massive importation of *bozales* did seem to have triggered some pidginization. Still, even the offspring of these pidgin speakers were monolingual in the local dialect of American Spanish, with some subtle transfer phenomena (Lipski, 2009: Chap. 4; López-Morales, 1980; Valdés-Bernal, 1994: 7, 2007b: 42). In other words, the presumed creole features that have been documented in the writings of African-born slaves and in literary imitations of their speech can be explained as the result of imperfect, adult language learning (Lipski, 2009: Chap. 9; López-Morales, 1980: 115-116, 1998: Chap. 5; Megenney, 1999: 280; Valdés-Bernal, 2007a: 11), language contacts with, mainly, Haitian and Jamaican Creole (Lipski, 1996: 41-43, 2009: Chap. 9) and/or transfer from a Bantu L1 background (Lipski, 2004: 122-123, 2009: 300).

Turning now to the emergence of the pluralized presentational *haber* construction, against the backdrop of the previous discussion, this variant may have arisen as a novel interdialectal compromise form. Since treating the presentational schema with *haber* as an intransitive construction is a common error in the second-language acquisition of Spanish, it may also have emerged in L2 varieties. In any case, as the pluralized presentational *haber* construction constitutes a simplification vis-à-vis the singular construction (Blas-Arroyo, 1995-1996: 178),[17] this variant would have been incorporated in the koine.[18] This is supported by the fact that the phenomenon is found on the Canary Islands and throughout the Latin American continent with similar constraints.[19] Additionally, this perspective on the emergence of *haber* pluralization is also supported by the quantitative data presented in the previous chapter. That is, following Tagliamonte's (2002) comparative method for tracing the characteristics of source varieties, the similarities between the constraint rankings presented in Chapter 6.2.5 suggest that Cuban, Dominican, and Puerto Rican Spanish share a common ancestor, in which the change was already operative and was restricted by the same set of cognitive constraints. The linguistic history of Latin America, in turn, suggests that this common ancestor was the Antillean koine.

As indicated earlier, any hypothesis concerning the rise of *haber* pluralization in Canarian and Latin American Spanish should also be able to account for the pluralization of Spanish presentational *haber* and Catalan presentational *haver-hi* (see example 24) in Eastern Spain.

(24) *Hi han* coses que no canvien (Internet, Blog, Barcelona, http://goo.gl/Kv5WM4).
'*There are*$_{Plur}$ things that do not change.'

Let us start by reviewing the data available on these two phenomena. In Catalan, the pluralization of presentational *haver-hi* has been documented in the Central[20] and Valencian[21] dialects (Blas-Arroyo, 1995-1996, 1999: Chap. 2; Ramos-Alfajarín, 1998: 56; Rigau, 1993: 45, 51). In addition, as was the case for Spanish *haber*, it seems that presentational *haver-hi* has always been an exceptional verb in Catalan, prone to sporadic form-function reanalysis. Still, Ramos-Alfajarín (2001) can only document a handful of cases before the nineteenth century. In contrast, his data suggest that in the

[17] In the sense that it realigns the syntax of presentational *haber* expressions with both the dominant trend in Spanish and the cognitive salience of their NP arguments.

[18] The fact that the koineization process retained both the singular and the pluralized presentational *haber* construction was probably due to statistical preemption, which ensured that native speakers of Spanish born in Spain experienced the pluralized forms as a deviation from established usage patterns. As a result, after the stabilization of the koine, the variation was associated to social types (see Trudgill, 1986: 110).

[19] See Chapter 2.2.

[20] The city of Barcelona and surrounding area.

[21] The city of Valencia and surrounding area.

second half of this century "the generalization of verb agreement in the spoken language must have been a phenomenon that had a strong incidence" (Ramos-Alfajarín, 2001: 139).[22] For the Spanish variety of Valencia, in turn, Blas-Arroyo (1995-1996: 191-192) finds that bilingual speakers with a Spanish L1 background pluralize *haber* far less frequently than their fellow community members with a Catalan L1 background. Monolingual speakers of Spanish pluralize even less often, especially when they are first-generation immigrants from other regions of Spain. Still, even within this group, the rates of *haber* pluralization do not drop below 30% (Blas-Arroyo, 1995-1996: 199).

As was the case for Spanish presentational *haber* on the Canary Islands and in Latin America, the fact that, in the nineteenth century, Catalan presentational *haver-hi* became involved in large-scale variation in Barcelona and Valencia is at least partially predictable from these cities' social histories. That is, in the course of that century, Barcelona and Valencia grew exponentially. This was due in part to the arrival of important contingents of immigrant workers from the Spanish interior, who were attracted by the booming textile industries (Cruz-Arroyo, 2000: 615-639). Like in the Spanish overseas colonies of the sixteenth century, this probably triggered the koineization of the different source dialects of Spanish, from which the pluralized presentational *haber* construction may have emerged. In turn, adult language learning of Catalan may have triggered the actuation of the reanalysis of presentational *haver-hi* in this language. Subsequently, the existence of a pluralized presentational *habere*-structure in both languages probably reinforced the tendency, as is evident from the fact that Catalan-dominant speakers pluralize Spanish presentational *haber* more often in Blas-Arroyo's (1995-1996) study.[23] In other words, although language contact may have played a more prominent role in the Catalan sphere of influence, as Blas-Arroyo (1995-1996, 1999: Chap. 2) argues, the key elements in the actuation of the reanalysis of Spanish presentational *haber* and Catalan presentational *haver-hi* appear to have been koineization and adult language learning, just like in the sixteenth-century Spanish colonies.

[22] In the original: "la generalització de la concordança en la llengua parlada degué ser un fenomen que tingué una forta incidència" (Ramos-Alfajarín, 2001: 139).

[23] Once conventionalized, the pluralized presentational *haber* and *haver-hi* constructions probably spread from Barcelona and Valencia to the varieties of Spanish and Catalan spoken in the surrounding areas (i.e., in the Central and Valencian Catalan dialect areas), for which these two cities function as innovative centers (see Labov, 2010: Part D; Trudgill, 1983: Chap. 2-3).

3. Summary

This chapter started out with a Cognitive Construction Grammar approach to the historical development of the singular presentational *haber* construction in Late Latin. Then, the literature on the history of Spanish *haber* pluralization was reviewed. Earlier historical linguistic research of the emergence of *haber* pluralization features three distinct lines of argumentation: those that favor the view that the alternation emerged in Late Latin, those that favor the view that it emerged as a Latin American innovation in the eighteenth century, and those that favor the view that it constitutes a recent pan-Hispanic innovation. Subsequently, I have presented an alternative approach. Although this account is rather speculative, it does accommodate the three fundamental facts about *haber* pluralization that were outlined here (historic distribution, geographic distribution, presence in Central and Valencian Catalan), while at the same time being fully compatible with the comparative sociolinguistic data presented in the previous chapter. To wrap up this dissertation, in the following chapter, an overview of the most important results will be presented and the research questions will be answered.

8

Wrapping up

Part B has presented a comprehensive analysis of *haber* pluralization in Caribbean Spanish, including its origins. This chapter, the last one of this dissertation, will review the results and highlight their implications for the phenomenon, for the nature of the constraints that condition it, for variationist sociolinguistics, and for Cognitive Construction Grammar. Particularly, section 1 will summarize the main findings. Against this background, section 2 will provide answers to the research questions and place these answers in a broader theoretical perspective.

1. Summary of the argument

Part A was concerned with the backgrounds of this study. In Chapter 2, the pluralization of presentational *haber* was introduced and the dialectological and sociolinguistic literature on the alternation was reviewed. The review of the dialectological literature showed that *haber* pluralization constitutes a widespread alternation, which appears in Canarian, Latin American, and some varieties of Peninsular Spanish. In turn, the overview of the sociolinguistic literature suggested that the phenomenon is sensitive to the properties of the NP (its reference, proportion of subject use, or stage-level/individual-level semantics), the absence/presence of negation, and the verb tense. For this factor, earlier investigations have shown that *haber* pluralization occurs frequently with the imperfect tense, the compound tenses, and with aspectual or modal auxiliaries. The verb is also pluralized more often in affirmative clauses and with human-reference NPs. Additionally, *haber* pluralization was shown to correlate rather consistently with lower- and middle-class membership in Venezuela. However, none of the earlier variationist studies offers an analysis that goes beyond describing the effect of these factor groups on the rates of *haber* pluralization.

In Chapter 3, Cognitive Construction Grammar was introduced. Crucially, this framework treats language as a structured inventory of form-function pairings, which, at least in the case of argument-structure constructions, symbolize ICMs. Against this background, the central hypothesis that was explored in this study proposes that *haber* pluralization is symptomatic of an ongoing change in the argument structure of the presentational *haber* construction: <**AdvP** *haber* **Subj**> is replacing <**AdvP** *haber* **Obj**>. Assuming this main hypothesis, four potential constraints on the variation were

identified: markedness of coding, statistical preemption, structural priming, and Labov's (2001) Principles of Linguistic Change. Before concluding Part A, Chapter 4 introduced the sample, the fieldwork methods, the corpus-building decisions, and the statistical toolkit of this study.

Subsequently, Part B focused on the analysis of the data. Chapter 5 presented a thorough description of the presentational *haber* constructions, which showed that, apart from their associations to social types and the syntactic function of their nominal arguments, the pluralized and the singular variant are completely identical. Particularly, this chapter argued that both the variants of the presentational construction with *haber* include a hearer-new nominal argument that fulfills a zero argument role and functions as the trajector of the clause. Both of the variants of the presentational construction with *haber* also include a profiled adverbial phrase, which sets up the mental space in which the construction localizes the referent of the nominal argument. Additionally, it was shown that under certain discourse conditions, both the nominal argument and the adverbial phrase can remain implicit.

Against this background, Chapter 6 presented a quantitative approach to *haber* pluralization. The results unveiled that the typical action-chain position of the noun's referent is the linguistic feature that subsumes the other factors that were tested in previous investigations. Particularly, Chapter 6.2.1 indicated that typical action-chain heads favor *haber* pluralization. In addition, when negation is present, speakers are less likely to select the pluralized presentational *haber* construction in San Juan. For Havana and Santo Domingo, in contrast, the absence/presence of negation did not rise above the significance threshold. Still, the fact that more potential agents are encoded more often with the variant of pluralized *haber* that has a subject suggests that markedness of coding is a cognitive constraint on the variation, as hypothesis 1 proposes.

For the verb tense, Chapter 6.2.2 showed that speakers pluralize *haber* less often in synthetic expressions with verb tenses that occurred mainly in the singular presentational *haber* construction before the actuation of the change.[1] In contrast, speakers pluralize *haber* more frequently with other types of expressions. This suggests that *haber* pluralization is constrained by statistical preemption, as is claimed by hypotheses 2a-c.

Additionally, Chapter 6.2.3 revealed that speakers are more likely to pluralize presentational *haber* in contexts following a pluralized presentational *haber* clause. Conversely, speakers are less likely to pluralize presentational *haber* when they have just used or processed a singular presentational *haber* expression. This supports the

[1] That is, with the present or preterit tense.

idea that, first, *haber* pluralization is subject to structural priming, as is argued by hypothesis 3 and, second, that presentational *haber* occurs in two argument-structure constructions, as is claimed by the main hypothesis.

The examination of the interaction between these cognitive constraints, in turn, suggested that for the three varieties, structural priming and the preference for unmarked coding incite speakers to extend the use of the pluralized presentational *haber* construction to the present and the preterit tense, whereas statistical preemption works against this. The comparative sociolinguistic analysis also revealed that the three cognitive factors essentially have the same effect in the three Caribbean varieties. Besides supporting hypothesis 9a, these results also corroborate that *haber* pluralization is not constrained by highly specific linguistic factors, but rather by the three general cognitive factors, as Chapter 3.3.2 claims.

Contrary to the similarity of the results obtained for the linguistic factor groups, the results for the social factor groups suggest that speakers of the three Caribbean varieties clearly establish different associations between the rates of *haber* pluralization and social types. Particularly, for Havana and the youngest generation of San Juan, pluralized *haber* correlates with lower social class. However, in Santo Domingo and for older Puerto Ricans, *haber* pluralization is a feature of middle-class speech. Pluralized *haber* also occurs more often in the speech of women in San Juan. Although these patterns of social covariation do not confirm hypotheses 5 and 6, they still suggest that in the three varieties under study, the variation constitutes an advanced ongoing language change from below. This supports hypothesis 9b.

Finally, Chapter 7 presented a Cognitive Construction Grammar approach to the emergence of the singular presentational *haber* construction in Late Latin. Then, an attempt at reconstructing the rise of the pluralized presentational *haber* construction was presented. Most importantly, Chapter 7.2 argued that *haber* pluralization emerged as a consequence of the koineization process that shaped the Canarian and Latin American dialects. A similar argument was presented for the Catalan and Spanish varieties of Barcelona and Valencia. Although rather speculative, this approach is able to account for four facts about *haber* pluralization, namely, its history, its geographic distribution, its presence in Central and Valencian Catalan, and the comparative sociolinguistic data presented in Chapter 6.

2. Conclusion

Let us return now to the research questions posited in Chapter 3.2 for the discussion of the results. For ease of reference, the questions are repeated here.

I. Cognitive factors in *haber* pluralization
   - What are the cognitive factors that constrain the pluralization of presentational *haber* in Caribbean Spanish?
   - How can these constraints be modeled in Cognitive Construction Grammar?

II. Social factors in *haber* pluralization
   - What is the social distribution of the pluralization of presentational *haber* in Havana, Santo Domingo, and San Juan?
   - Do these distributions justify the characterization of the phenomenon as a linguistic change from below?

III. Comparison of the Caribbean dialects
   - What are the differences (if any) between the social distributions and the effect of the cognitive constraints as they are observed in the varieties of Havana, Santo Domingo, and San Juan?
   - What do these results indicate about the emergence of *haber* pluralization and the nature of the constraints that condition it?

Concerning the first set of questions, the data presented in Chapter 6 suggest that the results of this and earlier investigations of *haber* pluralization can be reduced to three general cognitive factors that seem to be involved, and which reflect domain-independent characteristics of human cognition: markedness of coding, statistical preemption, and structural priming. Since these are among the general cognitive constraints that are posited on linguistic expression as a whole in Cognitive Construction Grammar, these factors can be modeled without any difficulties in this framework. Moreover, it is actually the framework that correctly predicts that these three cognitive factors will constrain *haber* pluralization.

The results arrived at in this study have some profound implications for variationist sociolinguistics. First, they show that cognitive-linguistic theory may contribute to identifying in an empirically and psychologically more adequate fashion the specific linguistic features to which variation is sensitive.[2] Second, this dissertation suggests a pathway for variationist analyses to go beyond the mere description of data and to reach the theoretic-linguistic goal of 'explanatory adequacy' (Chomsky, 1965: 26-27). That is, by stating the effect of specific linguistic environments as reflexes of general cognitive constraints on linguistic expression, which, in turn, reflect characteristics of

[2] See the discussion in Chapter 6.2.1.1, which led to the identification of the noun's typical action-chain position as the relevant factor connected with the noun.

human cognition, the analysis reported in this study does not only describe the variation ('descriptive adequacy'), it also explains why the data are the way they are ('explanatory adequacy').

For the second set of questions, the summary of the results has already shown that in the three Caribbean varieties, *haber* pluralization correlates with social class. In San Juan, the pluralization of *haber* also covaries with gender. Although these results do not pattern as was predicted by the hypotheses, the findings still support portraying the variation as an advanced ongoing language change from below, which has arrived at different stages in the three communities under investigation. This is also supported by the differences noted between the varieties for the frequency of <**AdvP** *hubieron* **Subj**> and the fact that the absence/presence of negation is only a significant factor in San Juan.

The fact that the results show that speakers use *haber* pluralization to position themselves in terms of social class (and gender) illustrates that the study of sociolinguistic variation inevitably leads to a question that is of central concern to functionalist linguistics, namely, why do speakers select construction X more often than construction Y to encode a particular ICM? As a functionalist theory, Cognitive Construction Grammar is able to incorporate the answer provided to this question by variationist sociolinguistics: to express nonpropositional/social meaning (e.g., Eckert, 2008; Labov, 2010: 372). This further illustrates the potential Cognitive Construction Grammar possesses for the study of language variation and change. For cognitive semantics, in turn, this finding implies that linguistic alternations do not necessarily serve to encode small conceptual-semantic contrasts in meaning, but may also serve to express social meaning.

Turning now to the third set of questions, in general terms, no striking dissimilarities were found between the three varieties as to the overall rate of *haber* pluralization, the cognitive factors that shape its usage, and their interaction. Rather, the differences between the three varieties appear to be situated in the differing patterns of social covariation. As noted in Chapter 7.2, this suggests that *haber* pluralization emerged as a form-function reanalysis in the koineization process that took place on the Antilles in the early colonial years.

In sum, the data and analyses presented in this dissertation appear to warrant the following general conclusions. First, regarding the phenomenon studied in this investigation, priming effects across specific verb forms suggest that *haber* occurs in two presentational constructions. Second, the effect that is produced by the independent variables that were tested in this dissertation was shown to be attributable to the same three general cognitive factors that may constrain any type of linguistic encoding (markedness of coding, statistical preemption, and structural priming) and

speakers' desire to position themselves in terms of social types. This principled explanation for the statistical patterns reported in this dissertation, in turn, shows that Cognitive Construction Grammar allows us to model syntactic change and the variation that is inherent to it for what they are: a competition within the system, constrained by general cognitive factors that allows speakers to position themselves within social categories.

This page intentionally left blank

Part C

Appendices

This page intentionally left blank

Appendix A

Social class questionnaire

A continuación figuran algunos factores que pueden contribuir al prestigio o apreciación social de que goza una persona. En tu opinión ¿cuáles de estos factores son más importantes que otros? Jerarquízalos (1= poco importante; 5= muy importante; sólo puedes utilizar cada número una vez).

Educación (p. ej., el tener algún diploma)	
Profesión (p.ej., taxista, cuentapropista, maestro, juez)	
Casa (p.ej., vivir en una casa cómoda, en un solar, en una microbrigada…)	

Appendix B

Interview schedule

1. Saludos y tiempo

- ¿Cómo estás? ¿Estás animado/a?
- ¿Últimamente está haciendo mucha calor, verdad? ¿Es normal o es excepcional?
- ¿Recuerdas el tiempo que hizo el año pasado por estas fechas?
- Dicen que está cambiando el clima ¿qué tú crees que va a pasar?

2. Barrio y casa

2.1 Barrio

- ¿En qué barrio o municipio vives?
    - ¿En dónde queda?
    - ¿Qué tipo de casas hay en tu calle?
    - ¿Me las puedes describir?
- ¿Cuántos años tú llevas ahí?
- ¿Por qué decidiste mudarte a esta zona?
    - **Si lleva toda la vida ahí:**
        - ¿Te gustaría mudar?
            - ¿Por qué, por el trabajo o hay otras razones ?
    - ¿Te mudaste ahí por el trabajo o había otras razones?
    - **Trabajo**:
        - ¿Este barrio queda más cerca de tu trabajo?
    - **Otras**:
        - ¿Cuáles?
- Te acuerdas de la mudanza?
    - ¿En qué año fue? ¿Qué edad tenías entonces?
    - ¿Qué tal los primeros contactos con los vecinos? ¿Fueron agradables?
    - ¿Había cosas a las que tuviste que acostumbrarte?
- ¿Te gusta vivir ahí? ¿O sea, es un barrio agradable?
    - ¿Hay ciertas cosas que no te gustan? Por ejemplo ruido, olores….
    - ¿Y al contrario, hay cosas que sí te gustan?

    - ¿Si te dieran la oportunidad de irte a vivir en otro lugar, en este país o en otro, dónde te gustaría vivir?
        - ¿Por qué?
    - ¿Cómo tú crees que sería vivir por allá?
- ¿Tienes la impresión de que ha cambiado el barrio en que vives durante el tiempo que llevas ahí?
    - ¿Qué cambios ha habido?
    - ¿Qué tú recuerdas de cuando eras pequeño?
    - ¿Cómo ves estos cambios, son para bien o para mal?¿cuéntame?
- Si el barrio ha cambiado, me imagino que en la ciudad entera habrá habido más cambios todavía, ¿verdad?
    - ¿Recuerdas cómo la cuidad era antes?
    - ¿Me podría nombrar 5 cosas que hay hoy y que no había cuando tú eras niño?
    - ¿La cuidad ha crecido? ¿Cuántos habitantes crees que podía haber durante tu niñez?
    - ¿Cuántos habrá ahora?
    - ¿Cuántos va a haber dentro de veinte años?
    - ¿Qué más cambios ha habido?
    - ¿Cómo ves estos cambios, son para bien o para mal?
        - ¿Por qué?
        - ¿Cómo te hubiera gustado que esta ciudad fuera?
        - ¿Cómo tú ves la delincuencia?
        - ¿Crees que antes había más delincuentes o menos?
        - ¿Hay muchos lugares peligrosos en esta ciudad?
            - ¿Me puedes enumerar los cinco lugares más peligrosos?
            - ¿Y estos son peligrosos todo el tiempo, o sólo de noche?
        - ¿Y en tu barrio, hay mucha delincuencia?
        - ¿Qué tú crees que se puede o debe hacer para acabar con ella?
        - ¿Has oído hablar de algún delito en tu barrio?
        - ¿Qué pasó?
- En tu opinión, ¿qué cambios va a haber en este ciudad en el futuro?
    - ¿Por qué tú piensas esto?
- ¿Tú crees que dentro de veinte años habrá más o menos pobres en esta ciudad?
    - ¿Por qué?
- ¿Conoces a mucha gente de tu barrio?
    - ¿Qué tal te llevas con ellos?
    - ¿Cómo son tus vecinos?
    - ¿Son como te gustaría que fueran?
    - ¿Qué relación tienes con ellos?

        - ¿Hacen cosas juntos?
    - ¿Qué relación te gustaría tener con ellos?
    - ¿Qué tú haces para mantener o mejorar esta relación?
- Hay quienes dicen que hoy en día todos estamos tan ocupados que no tomamos ya el tiempo para hablar con la gente en la calle. ¿Qué tú piensas?
    - ¿En la calle en que vives/en esta calle, los vecinos todavía se hablan cuando se cruzan?
    - Qué tú recuerdes, ¿cuándo tú eras niño/a, había más conversaciones en la calle?
        - ¿De qué la gente hablaba?
        - **Si ahora menos:**
            - ¿Piensas que es una pena que esta costumbre se haya perdido?
            - ¿En tu opinión por qué esto sucedió?
- ¿Hay vecinos que conoces lo suficientemente bien como para ir a visitarlos sin avisar?
    - ¿Quienes?
    - ¿En dónde viven?
    - ¿Los visitas mucho?
    - ¿Hacen cosas juntos? Por ejemplo, ir de compras, a la playa….
- ¿Los vecinos u otros conocidos del barrio te invitarían a tomar café si te toparas con ellos en la calle?
    - ¿Quién?
    - ¿Y eso pasa mucho?
- ¿Hay gente del barrio o del barrio en que te criaste con la que te gustaría pasar más tiempo?
    - ¿Por qué no se ven tanto ya?
    - ¿Hay razones especiales para ello?
- ¿Alguna vez te peleaste o te discutiste con un vecino u otra persona del barrio?
    - ¿Por qué fue?
    - ¿Se reconciliaron pronto o estuvieron un tiempo sin hablar?
    - ¿Cómo solucionaron su problema?
    - ¿Y ahora se llevan bien?
    - ¿Eres una persona rencorosa o perdonas fácilmente?
- ¿Hay lugares en tu barrio en donde la gente se puede encontrar?
    - **No**:
        - ¿Tienes alguna idea de por qué no los hay?
        - ¿Qué tú recuerdes, los hubo antes?
    - **Sí:**
        - ¿Cómo las gentes le decían?

        - ¿Por qué no los hay ya?
        - ¿Qué tipo de actividades había ahí?
        - ¿Iba mucha gente?
    - **Sí**:
        - ¿Cómo le dicen?
        - ¿Que tú recuerdes, siempre los ha habido?
        - ¿Qué tipo de actividades hay ahí?
        - ¿Vas mucho?
        - ¿Antes iba más gente?
- En tu barrio, ¿Hay organizaciones comunitarias?
    - **Sí**:
        - ¿Qué hacen? ¿Qué tipo de actividades organizan?
            - ¿Vas a estas actividades?
            - ¿Por qué?
    - **No**:
        - ¿Hubo organizaciones de este tipo antes?
            - **Sí**:
                - ¿Qué hacían?
                - ¿Por qué no existen ya?
                - ¿Piensas que es una pena?
- ¿En tu barrio, hay algún bar, colmado o cafetería a dónde vas mucho?
    - ¿Me explicas cómo se llega ahí?
    - ¿Me lo puedes describir?
    - ¿Qué sueles tomar ahí?
- ¿Hay buenas panaderías en tu barrio?
    - ¿También hacen bizcochos y dulces?
- ¿Donde se consigue el mejor X del barrio?

2.2 Casa

- ¿Hablando de tu casa, me la puedes describir?
    - ¿Es grande o chica?
    - ¿Cómo está dividida?
    - ¿Cuántos cuartos hay?
    - ¿Cuántos baños hay?
    - ¿Tienes aire acondicionado o ventiladores?
    - ¿Tienes televisión?
    - ¿Tienes computadora?
    - ¿Tu casa tiene algún defecto, o sea hongos, grietas, fugas…?
    - ¿Hiciste reformas o estaba igual cuando te mudaste ahí?
        - ¿Cómo era antes?

    - ¿Hay cosas que le quieres cambiar, o sea por ejemplo pintar las paredes, cambiar el piso?
    - ¿Ya sabes cuándo lo vas a hacer?
- ¿Me puedes describir la casa en que te criaste?
    - ¿Era grande?
    - ¿Cómo estaba dividida?
    - ¿Cuántos cuartos había?
    - ¿Cuántos baños había?
    - ¿Había aire acondicionado o ventiladores?
    - **Primera generación:**
        - ¿Tenían televisión?
        - ¿Tenían computadora?
    - **Segunda generación:**
        - ¿Tenían televisión y radio?
    - ¿Y ésta tenía algún defecto?
    - ¿Tus papás le hicieron reformas?
        - ¿Cuáles?
- Si comparas tu casa paterna con la tuya, ¿cuál sería la más cómoda?
    - ¿Por qué?
- ¿Qué tal estaba tu cuarto?
    - ¿Me lo describes?
    - ¿Lo compartías con otra persona?
        - ¿Con quién?
        - ¿No te hubiera gustado tender tu propia habitación?
    - ¿Dónde jugabas?
        - ¿En la sala o en el cuarto?
        - ¿Por alguna razón especial o simplemente fue así?

3. Escuela

- ¿Te acuerdas de la escuela primaria? ¿De la secundaria/high?
    - ¿Fue una escuela pública o privada?
    - ¿Había monjas o monjes?
- ¿Recuerdas el edificio?
    - ¿Me lo describes?
    - ¿Todavía existe?
        - ¿En dónde está? ¿En qué barrio?
    - ¿En los salones, había aire acondicionado o solamente había abanicos?
    - ¿Cuántos alumnos habría en tu época?
    - ¿Había niños y niñas o solamente chicos/chicas?

- ¿Cuántos habrá ahora?
- ¿Ustedes tenían que ponerse uniforme?
    - **Sí:**
        - ¿Me lo podrías describir?
        - ¿Había castigos para los que no llevaban el uniforme?
        - ¿Se permitía llevar pantallas (PR)/aretes?
        - ¿Había cortes que quedaban prohibidos?
    - **No:**
        - ¿Que te ponías entonces?
        - ¿Había trajes que quedaban prohibidos?
        - ¿Y cortes?
        - ¿Se permitía llevar pantallas/aretes?
        - ¿Había castigos por no vestirse conforme con las reglas?
- ¿Qué tal estaban los maestros?
    - ¿Había algunos muy malos?
    - ¿Había también algunos muy buenos?
    - ¿Piensas que había más buenos maestros que ahora?
        - ¿Por qué?
        - ¿En tu opinión por qué eso ha cambiado?
    - ¿Había cosas que les enfadaban? ¿Cuáles?
    - ¿Qué castigos había?
    - ¿Alguna vez te regañaron o te castigaron por algo que no hiciste?
        - ¿Cómo esto te hizo sentir?
        - ¿Se lo dijiste a tus papás?
        - ¿Cómo reaccionaron? ¿Te creyeron?
        - ¿Fueron a la escuela a hablar con el maestro?
- ¿Los maestros te daban muchas tareas?
    - ¿Para qué materias?
    - ¿Recuerdas cuánto tiempo las tareas te tomaban?
- ¿Había materias que te gustaban más que otras?
    - ¿Cuál te gustaba más?
    - ¿Por qué
    - ¿Cuál te gustaba menos?
    - ¿Por qué
- Pasaban notitas en la clase?
    - **Sí**:
        - ¿De qué trataban?
    - **Sí**:
        - ¿Alguna vez el maestro te pilló?
            - **Sí:** ¿Qué pasó entonces?

- ¿Hacían excursiones?
    - ¿A dónde fueron?
    - ¿Algo interesante ocurrió?
- ¿Te acuerdas de lo que solían hacer durante el recreo?
    - ¿Jugaban juegos?
        - ¿cuáles?
    - ¿Me puedes explicar las reglas?
- ¿Todavía conservas a amigos de la primaria?
    - ¿De la high/secundaria ?
    - ¿Se organizan reuniones?
        - **Sí:**
            - ¿Cuándo es la próxima?
            - ¿Vas?
            - ¿Por qué?
            - **No va**:
                - Si fueran diferente de como son, ¿irías?
        - **No:**
            - ¿Si se organizaran, irías?
            - ¿Hay razones especiales para ello?
            - ¿Si te tocara organizar una reunión, cómo harías?
- ¿Y después de la primaria/de la secundaria/high, hasta qué grado llegaste?
    - ¿Obtuviste el diploma?
    - ¿Hasta dónde tus papás querían que fueras?
    - ¿Y hasta dónde ellos llegaron en la escuela?
    - **Si tiene formación universitaria:**
        - ¿Estudiaste en el extranjero?
        - ¿Adónde fuiste?
        - ¿Por cuánto tiempo?
        - ¿Volvías de vez en cuando?
        - ¿Cuándo?
        - ¿Cómo fue esta experiencia para ti?
    - **Si no obtuvo el diploma:**
        - ¿Por qué dejaste los estudios?
        - ¿Qué tus papás dijeron de ello?

4. Trabajo

- Después de dejar los estudios/después de graduarte, ¿fue fácil encontrar trabajo?
    - ¿Cuál fue tu primer trabajo? ¿En qué empresa fue?
    - ¿Qué tenías que hacer?
    - ¿Qué tal tu contacto con los colegas?
    - ¿Había más colegas de tu generación?
    - ¿Te acuerdas de cuánto ganabas?
    - ¿Recuerdas qué hiciste con tu primer salario?
    - ¿Hubo cosas que realmente quisiste comprar con este dinero?
- ¿Qué profesión tus papás querían que escogieras?
    - ¿Por qué?
    - ¿Había profesiones o estudios que ellos veían mejor?
    - ¿Estaban contentos con el puesto que conseguiste?
    - ¿Entonces, en casa, (no) hubo discusiones a este respecto?
    - ¿Habría profesiones que tus padres nunca habrían aceptado?
        - ¿Cuáles?
        - ¿Por qué?
- **Primera generación:**
    - ¿Hablando de tus papás,
        - ¿Qué tipo de personas son, o sea, en cuanto a carácter?
        - ¿Cómo se ven?
        - ¿Te pareces más a tu madre o a tu padre?
    - ¿Ya están jubilados?
        - **Sí:**
            - ¿Desde hace cuánto?
        - **No:**
            - ¿Cuánto les falta?
        - ¿A qué se dedican?/se dedicaban?
        - ¿En qué consiste/consistía su trabajo?
        - ¿Y tu madre trabaja/trabajaba también?
        - ¿Que tú recuerdes, siempre se han dedicado/dedicaron a eso? ¿O hubo tiempos que hacían otra cosa?
            - **Sí:**
                - ¿A qué se dedicaron entonces?
- **Segunda generación:**
    - ¿Siguen vivos y sanos?
        - ¿Qué tipo de personas son/fueron, o sea, en cuanto a carácter?
        - ¿Cómo se ven/se veían?
    - ¿A qué ellos se dedicaban?

        - ¿En qué consistía su trabajo?
        - ¿Dónde trabajaba su padre?
        - ¿Y su madre trabajaba también?
        - ¿Que usted recuerde, siempre se dedicaron a eso? ¿O hubo tiempos que hacían otra cosa?
            - **Sí:** ¿A qué se dedicaron entonces?
- ¿A qué te dedicas ahora?
    - ¿Te gusta?
    - ¿Qué aspectos te gustan más?
    - ¿Qué aspectos te gustan menos?
    - ¿Eso es lo que siempre has querido hacer?
        - **Si no**: ¿Por qué no lo hiciste?
            - ¿Cómo te imaginas tu vida si hubieras sido…?
    - ¿Qué sueles hacer en un día normal, laborable? O sea a qué hora te levantas, a qué hora te vas al trabajo, etc., etc.

5. Solidaridad

- ¿En la calle o el barrio en que vives, hay gente a los que puedes pedir que te echen la mano?
    - ¿Quiénes son?
    - ¿Cuál es tu relación con ellos?
- ¿Tú ayudas a otros? Pienso por ejemplo en los vecinos…
- ¿Si te cayeras enfermo/a y necesitaras ayuda, en tu barrio habría gente que te ayudaría?
    - ¿Tendrías que pedírselo o te lo ofrecerían?
- Y si, por ejemplo, estuvieras cocinando un bizcocho y te faltaran huevos, los vecinos te los prestarían?
    - ¿Y hay cosas que tú les prestas a los vecinos?
    - ¿Cuáles?

6. Cenas familiares/cultura culinaria

- ¿Cuando eras niño/a quién te cocinaba?
    - ¿Y X era buen cocinero/a?
    - ¿A veces lo/a ayudabas?
        - **Sí**: [¿Qué cocinaban juntos?
            - ¿Cómo se prepara?]
- En muchos países del mundo es costumbre comer mejor los domingos. Cuando eras niño/a, ¿había platos que sólo se preparaban los domingos?

    - ¿Cuáles?
    - ¿Cómo se elabora?
    - ¿Y ahora, en casa continúan esta tradición?
- ¿Te acuerdas de algún plato especial o muy rico que tu X te hacía o todavía te hace?
    - ¿Cuál era?
- Había platos que no te gustaban para nada?
    - ¿Cuéntame?
    - ¿Por qué X no te gustaba?
    - ¿Tus papás te obligaban a terminar lo que había en tu plato?
    - ¿Y ahora X ya te gusta un poco?
- Y al contario, ¿había platos que a ti te gustaban mucho pero no a los demás?
    - Cuéntame.
- ¿Te gusta cocinar?
- **Sí:**
    - Eres buen cocinero/buena cocinara?
    - ¿Hay ciertos platos de que la gente dice "mira él/ella sabe hacerlo muy, pero muy bien"?
        - ¿Cuáles?
        - ¿En tu opinión, por qué tu manera de hacerlo les agrada tanto a los demás?
        - ¿Compartes la receta o es tu secreto?
    - ¿Cuál es tu plato favorito?
        - ¿Cómo se prepara?
        - ¿Quién te enseño esta receta?
        - ¿Hay gente que lo hace de otra manera?
    - ¿Cuál es la comida típica de X?
    - Cómo se prepara?
        - ¿Por qué se añade X?
        - ¿Cuándo se le echa?
        - ¿Habrá gente que lo hace de otra manera?
    - ¿En este país, hay diferencias regionales en cuanto a comida?

- **No**:
    - ¿Pero sabes cocinar? ¿O sea, te defiendes?
    - ¿Qué platos sabes hacer?
        - ¿Cómo tú lo preparas?
        - ¿Por qué se añade X?
        - ¿Cuándo se le echa?
        - ¿Quién te enseñó esta receta?

  - ¿Hay platos que te gustaría aprender cómo cocinarlos?
    - ¿Cuáles?
    - ¿Por qué?
  - ¿Qué crees que es necesario para uno ser un buen cocinero?
  - ¿En este país, hay diferencias regionales en cuanto a comida?

7. Papás y familia

- ¿Cómo tus papás te criaron?
  - ¿Fueron severos contigo?
    - ¿Que tú recuerdes, había más padres como los tuyos?
  - ¿Cuando eras chico/a, qué más te gustaba de ellos?
- ¿Tienes la impresión de que los papás de hoy día son más estrictos o es al revés?
  - ¿Por qué?
  - ¿Qué indicios hay?
- ¿Qué tipo de hijo/a eras? ¿Eras dulce, malo/a, llorón/a?
  - ¿Piensas que los hijos de hoy son diferentes?
  - ¿Qué cambios ha habido a este respecto?
  - ¿En tu opinión por qué es así?
- ¿Cuándo eras niño, qué pasatiempos había?
  - ¿Cuál era tu pasatiempo favorito?
  - ¿Piensas que los niños de hoy todavía lo hacen?
  - ¿Piensas que los niños de hoy tienen menos imaginación que ustedes a esta edad?
    - ¿Por qué?
    - ¿Cómo se ve?
- ¿Hacías travesuras? Cuéntame
  - ¿Te castigaban por ellas?
  - ¿Cómo te castigaban?
  - ¿Te castigaba tu mama o eso lo hacía tu papá?
  - ¿Había más cosas por las que te castigaron?
- ¿Alguna vez tus papás te echaron la culpa de algo que no hiciste?
  - ¿Por qué fue?
  - ¿Sólo te regañaron o también te castigaron?
  - ¿Cómo eso te hizo sentir? ¿Te dio mucha rabia?
  - ¿Después ellos se percataron de su error?
  - ¿Y qué pasó entonces?
- ¿Cuando eras adolescente, tus papás te permitían salir con amigos?
  - ¿Te daban dinero para ello?

    - ¿Cuánto de daban?
    - ¿Te alcanzaba?
    - ¿Qué se podía comprar con ese dinero?
    - ¿Al salir, te imponían horario?
    - ¿Alguna vez llegaste tarde?
        - ¿Qué pasó entonces?
- ¿Cuando salías con los amigos, qué te ponías?
- ¿A dónde salían? ¿A bailes, a casa de amigos?
    - ¿Cómo llegaban hasta ahí? ¿Iban a pie, en guagua en carro o tus papás te llevaban?
    - ¿Cómo eran estas salidas/bailes?
    - ¿Qué música se ponía ahí?
    - ¿Bailabas?
    - ¿A veces había peleas en donde estaban?
- ¿Cuando había problemas entre tú y tus padres, con quién hablabas para resolverlos?
    - ¿Por qué?
    - ¿Fue porque X era más permisivo/a?
- ¿Tienes hermanos?
    - ¿Cuántos?
    - ¿Y tú eres el/la mayor, menor, mediano?
    - ¿Cuánto ellos te llevan/Cuánto tú les llevas?
- ¿Tus hermanos y tú se llevaban bien o se peleaban mucho?
    - ¿Por qué se peleaban?
- Y ahora se llevan bien?
- ¿Alguna vez les jugaste una broma a tus hermanos?
- ¿Hubo momentos en que te avergonzaste de tu familia, ya sean tus hermanos, papás, abuelos…?

8. Amigos

- ¿Cuando tú eras chico, pasabas mucho tiempo con los amiguitos/las amiguitas.
    - ¿Todavía los ves de vez en cuando?
        - No: ¿Cómo fue que se perdieron de vista?
    - ¿Sabes dónde viven?
- ¿Son iguales los amigos que uno tiene de niño a los que uno tiene de adulto?
    - ¿Qué diferencias hay?
    - ¿En tu opinión, por qué es así?
- ¿Qué es para ti un amigo?
- ¿Pasas mucho tiempo con los amigos?

- ¿Qué tipo de actividades hacen juntos?
- ¿Hay actividades que tú puedes hacer con los amigos, pero no con la familia?
  - ¿Cuáles?
- ¿Qué hacen cuando salen?
  - ¿Tú y los amigos, van al cine?
    - ¿Cuál fue la última película que llegaste a ver?
      - ¿De qué trata?
      - ¿Te gustó?
      - ¿Por qué
      - ¿Cuál es tu película favorita?
      - ¿De qué trata?
      - ¿Por qué te gusta tanto?
  - ¿Los amigos/as y tú salen de noche a bailar?
    - ¿Qué música se pone ahí?
      - ¿Toca una banda o un dj?
      - ¿Ahora el reggaetón está muy de moda, qué piensas de este género de música?
      - ¿Te gusta? ¿No te gusta?
        - ¿Por qué?
        - **Si le gusta**:
          - ¿Hay cantantes que escuchas más que otros?
          - ¿Qué te gusta más : el reggaetón de Puerto Rico o el Cubatón/el reggaetón que se hace por acá?
          - ¿Por qué?
  - ¿Qué [otra] música te gusta?
    - ¿Hay cantantes que escuchas más que otros?
      - ¿Por qué?
      - ¿Cuál es tu cantante favorito?
    - ¿Van a conciertos?
      - ¿Cuáles son los conciertos que llegaron a ver?

9. Sentido común

- ¿En tu opinión ¿qué es sentido común?
    - ¿Hay ciertas decisiones que van en contra del sentido común?
    - ¿Cuáles?
- ¿Entre tus conocidos, hay algunos que tienen mucho o muy poco sentido común?
    - ¿Cómo se ve eso?
- ¿Piensas que el sentido común aumenta con los años?
    - ¿Por qué?

10. Juegos

- ¿Cuando eras niño, qué juegos jugabas después de la escuela?
    - ¿Me puedes explicar las reglas?
    - ¿Con quién jugabas?
    - ¿Jugaban al escondite?
    - ¿Como decidían a quién le tocaba buscar a los demás?
- ¿Había juegos que les gustaban más a las niñas que a los niños?
- ¿Qué juguetes tenías?
    - ¿Había juguetes que eran especiales para ti?
    - ¿Por qué?
    - ¿Quién te los regaló?
    - ¿En qué ocasión?
    - ¿Qué más hacías después de las clases? ¿Hacías deporte?
    - ¿Qué deporte?
    - ¿Me puedes explicar las reglas?
        - ¿Competías? ¿Ganaste algunos juegos?
        - ¿Qué pasó?
    - ¿Había deportes que a ti te habrían gustado hacer, pero tus papás no te dejaron?
        - ¿Cuáles?
        - ¿Por qué no te lo dejaron hacer?
- Tenías hobbies?
    - ¿Cuáles?
    - ¿Por qué empezaste a hacerlo?
    - ¿Todavía lo haces?
        - **Sí:**
            - ¿Te toma mucho de tu tiempo?
        - **No:**
            - ¿Por qué lo dejaste?

        - ¿Fue por los estudios, el trabajo u otra razón?
- Y los adultos, jugaban juegos?
    - ¿Cuáles?
    - ¿Me puedes explicar las reglas?
    - ¿Jugaban con las cartas?
        - ¿Qué juegos había?
        - ¿Me puedes explicar las reglas?
    - ¿Jugaban dominó?
    - Ahora, ¿Tú también haces estos juegos?
        - ¿Te gustan?

11. Cumpleaños

- ¿Cuándo tú cumples?
    - ¿Esta fecha tiene sus desventajas?
    - ¿Cómo celebras tu cumple?
- ¿Que te recuerdes, cuál fue la mejor fiesta de cumpleaños que tuviste?
    - ¿Cuéntame?
- ¿Cuál fue la peor?
    - ¿Cuéntame.
- ¿Cuál fue la mejor a la que tú pudiste asistir?
- ¿Cuando eras niño, había veces que los amiguitos/las amiguitas se quedaban a dormir en tu cumpleaños?
    - ¿En otra ocasión?
    - ¿Tus papás los regañaban cuando no querían dormir?
    - ¿Qué hacían entonces?
- ¿Cuál fue la mejor de estas fiestas de pijamas?

12. Turismo interno

- ¿Qué lugares de la isla visitaste?
    - ¿A dónde fuiste?
    - ¿Por cuánto tiempo?
    - ¿Con quién?
- ¿Los hay que te gustaría volver a visitar?
- ¿Cuáles me puedes aconsejar?
- ¿Los domingos van a la playa?
    - ¿Qué playa prefieres?
    - ¿En dónde queda?
    - ¿Cómo se llega?

    - ¿Me la puedes describir?
- ¿Cuando haces excursiones, cómo vas? ¿En carro o coges un carro público o una guagua, tren, vuelo interno?
    - ¿Hablando de carros, tienes carro?
    - ¿Qué marca es?
    - ¿Me lo puedes describir?
    - ¿Cuán importante es para ti guiar/manejar un carro como el que tienes?
- ¿Alguna vez tuviste problemas viajando, o sea no pudiste volver porque se te dañó el carro, el parador/la casa estaba completo/a…?
    - ¿Qué hicieron entonces?
    - ¿Cuándo fue?
    - ¿Ya hubo celulares en ese tiempo?
    - ¿Cómo solucionaron el problema?
- ¿Hay lugares de la isla a donde nunca has podido ir, pero que te gustaría visitar?
    - ¿Cuáles?
    - ¿Me puedes contar más sobre ellos?
    - ¿Por qué quieres ir?
    - ¿Ya tienes idea de cuándo vas a visitarlos?

13. Tradiciones, costumbres y fiestas

- ¿Qué tradiciones hay?
    - ¿Las continúas?
    - ¿Hay todavía muchas familias que las continúan?
        - ¿En tu opinión, por qué es así?
- ¿Qué fiestas celebran (por ejemplo, Navidades, Año Nuevo, …)?
    - ¿Cómo las celebran?
    - ¿Qué ropa te pones en estas fiestas?
    - ¿Qué suelen comer?
        - ¿Me puedes explicar cómo se preparan, o de qué se hacen?
            - ¿Estos platos sólo se comen en esos momentos o hay más ocasiones en que se preparan?
            - ¿Cuáles?
        - ¿Cuál es la comida típica de navidad?
            - ¿Cómo se hace?
            - ¿Habrá familias que comen otra cosa?
            - ¿Siempre ha sido así o ha habido cambios a este respecto?
    - ¿En algunas de estas fiestas, se dan regalos?
        - ¿Cuándo?
        - ¿Como los abres, cuidadosamente o rasgas el papel por completo?

        - ¿Alguna vez recibiste un regalo que no te gustaba para nada?
            - ¿Cómo reaccionaste?
            - ¿Se lo dijiste al que te lo regaló?
            - ¿Cómo él/ella reaccionó?
- ¿Cuál fue la mejor de estas fiestas familiares?
    - ¿Por qué?
- ¿Cuál fue la peor?
    - ¿Por qué?
- ¿Hay otras tradiciones por acá, este, fiestas patronales, carnavales?
    - ¿Antes sí los había?
    - ¿Ha habido cambios a este respecto?
    - ¿Por qué?
- ¿Qué sueles hacer en Nochevieja?
    - ¿La celebras con amigos o con la familia?
    - ¿Hay platos que sólo se comen esta noche?
        - ¿Cómo se preparan, o de qué se hacen?

14. Experiencias poco comunes

- ¿Alguna vez te pasó algo extraño, o sea, algo que tú no puedes explicar?
    - ¿Cuándo fue? ¿Qué pasó?
    - ¿Volvió a pasar después?
    - ¿Cómo lo explicarías?
- En algunas familias, hay una persona que sabe como predecir el futuro.
    - ¿Hay una persona así en tu familia?
    - ¿Le hacen caso?
    - ¿Te acuerdas de algún acontecimiento que él/ella predijo?
- ¿Eres supersticioso?
    - ¿Qué haces para tener buena suerte?
    - ¿Tienes suerte con las cartas?
    - ¿Y con el dominó?
    - ¿Y con las mujeres/los hombres?
- ¿Alguna vez presenciaste un accidente, un tiroteo o algún desastre?
    - ¿Qué pasó?
    - ¿Si pudieras volver al pasado, habría cosas que harías de otra manera?
    - ¿Cuáles? ¿Y por qué?
- ¿Alguna vez estuviste ingresado en el hospital por una urgencia?
    - ¿Cuál fue la emergencia?
    - ¿Te quedaste mucho rato?
- ¿Alguna vez estuviste en peligro de muerte?

  - ¿Qué pasó?
  - ¿Cómo te salvaste?
  - ¿Qué hubiera pasado si...?
  - ¿Y si...?
- ¿Te acuerdas de otro momento específico en que tuvieras mucho miedo?
  - ¿Cuando fue?
  - ¿Qué pasó?
  - ¿Cómo te sentiste después, cuando estabas a salvo?
- ¿Conoces a gente que no le tiene miedo a nada?
  - ¿Realmente no tienen miedo, o es que no quieren admitir que sí lo tienen?

15. Eventos históricos importantes

- ¿Te acuerdas del huracán X?
  - Yo nunca he vivido un huracán, ¿me puedes contar cómo es?
  - ¿Qué se hace por acá para evitar los daños y los peligros?
  - ¿Como ese huracán les afectó a ti y tu familia?
  - ¿Crees que hubo muertos en esta ciudad? ¿Cómo cuántos?
  - ¿Hubo muchos daños?
  - ¿Cómo se organizaron para repararlos?
  - ¿Ha habido más huracanes como éste que tú recuerdes?
  - ¿Crees que en un futuro próximo va a haber más huracanes como éstas?

16. Lengua

- Cuando tú escuchas la gente hablando por acá, ¿oyes algunas particularidades? ¿O sea en relación con la manera como pronuncian los sonidos, las palabras que emplean?
- ¿Se dice que la lengua está cambiando. ¿ Tienes esta impresión, o sea, piensas que tus hijos o tus nietos van a hablar de otra manera?
  - ¿Cómo tú ves estos cambios? ¿Son para bien o para mal?
  - **Segunda generación:**
    - ¿Oyes diferencias entre tu manera de hablar y el habla de los jóvenes?
    - ¿Hay ciertos rasgos en el habla de los jóvenes que no te gustan?

  - **Primera generación:**
    - ¿Oyes algunas diferencias entre tu manera de hablar y la de, por ejemplo, tu abuelo?

        - Se dice que la lengua está cambiando. ¿Tienes esta impresión, o sea, piensas que tus hijos o tus nietos van a hablar de otra manera?
- Cuando tú escuchas a un hablante del español ¿sabes de dónde es?
    - ¿Cómo lo sabes? ¿Qué indicios hay?
    - ¿Hay acentos del español que te suenan mejor que otras?
    - ¿Qué dialecto del español te gusta más?
    - ¿Piensas que el tuyo suena mejor, peor?
    - ¿Qué te parece el español cubano/el español dominicano/el español puertorriqueño/el español de España?
        - ¿Piensas que suena mejor, peor que el tuyo?
        - ¿Por qué?
        - ¿En qué te hace pensar?
    - ¿Oyes diferencias entre el acento de X y el de X?
        - ¿Qué diferencias hay?
    - ¿En tu opinión, hay mucha influencia de otras lenguas sobre el español que se habla por acá?
        - ¿De qué lenguas?
        - ¿Cómo se ve?
- ¿Tú cómo tratas a tus amigos de 'tú' o de 'usted'?
    - ¿Y si son personas mayores?
    - ¿Y si son desconocidos?
    - ¿Cómo te sientes si una persona te trata de 'usted'/de 'tú'?
    - ¿Y si es una persona mayor, hombre o mujer, al que por ejemplo le preguntas por una calle?
    - ¿Y a tu médico?
    - ¿Y a extranjeros?
    - ¿Y cómo te gusta que te traten a ti?
    - Si una persona más joven te trata de 'tú' qué te parece?
    - ¿Todavía se emplea el trato de 'su merced'?
    - ¿En qué ocasiones se emplea?
    - ¿Qué forma te parece más respetuoso: el usted o el su merced?
    - ¿En tu opinión hay diferencias entre los dos?

Appendix C

Story-reading task

The reading task was adapted from the Internet (http://goo.gl/8W61jT). The original story was conserved entirely, but the syntax was updated to reflect a less archaic type of language. I also inserted the selection contexts.

Juan Sin Miedo

En una pequeña aldea, **había/habían** (1) un anciano padre y sus dos hijos. El mayor era trabajador y llenaba de alegría el corazón de su padre, mientras el más joven sólo le daba disgustos. Un día el padre lo llamó y le dijo:

— "Hijo mío, sabes que no **hay/hain** (2) muchas cosas que yo pueda dejarles a tu hermano y a ti, y sin embargo tú aún no **aprendiste/aprendías** (3) ningún oficio que te sirva para ganarte el pan. ¿Qué te gustaría aprender?"

Y le contestó Juan:

— "Bueno, **hay/hain** (4) varias cosas de que me gustaría saber cómo hacerlas. Muchas veces yo oigo relatos en que hay monstruos, fantasmas, fieras y al contrario de la gente, no siento miedo. Papá, yo quiero aprender a tener miedo."

El padre, enfadado, le gritó:

— "Estoy hablando de tu futuro, y ¿tú, tú quieres aprender a tener miedo? Si es eso lo que quieres hacer, pues márchate a aprenderlo. Espero que en el camino **haya/hayan** (5) varias situaciones que te **inspiran/inspiren** (6) miedo."

Juan recogió sus cosas, se despidió de su hermano y de su padre, y emprendió su camino.

Cerca de un molino encontró a un sacristán con quien se puso a hablar. El joven se presentó como Juan Sin Miedo.

— "¿Juan Sin Miedo? ¡Extraño nombre!" – El sacristán se **admiró/admiraba**. (7)

Juan dijo:

— "Ya vas a ver, no **hay/hain** (8) peligros, ogros, fieras, bestias que me den miedo, porque nunca de mi vida yo he conocido el miedo. Partí de mi casa para conocer lo

que es, pero hasta el momento en el camino no **hay/hain** (9) personas, no **hay/hain** (10) situaciones, no **hay/hain** (11) animales que me inspiren miedo. Sí que ayer, **hubo/hubieron** (12) dos lobos que querían devorarme, anteayer **hubo/hubieron** (13) unos ladrones que trataban de matarme y **ha habido/han habido** (14) dos veces que yo tenía que brincar un abismo de treinta pies de ancho y todo esto fue muy molesto, pero miedo como tal no tuve".

El sacristán dice:

— "Quizá yo pueda ayudarte. Cuentan que más allá del valle, muy lejos, hay un castillo encantado por un mago. El rey que allí gobierna prometió la mano de su linda hija a aquel que **consigue/consiga** (15) recuperar el castillo y el tesoro. Hasta ahora, todos los que lo intentaron huyeron asustados o murieron de miedo".

Juan se animó:

— "Quizá, quizá allí **haya/hayan** (16) los peligros necesarios para yo sentir el miedo".

Juan decidió caminar, vio a lo lejos las torres más altas de un castillo en el que no **había/habían** (17) banderas. Se acercó y se dirigió a la residencia del rey. Dos guardias reales cuidaban la puerta principal. Juan se acercó y **decía/dijo** (18):

— "Soy Juan Sin Miedo, y deseo ver a su Rey. Quizá él me permita entrar en su castillo y sentir a lo que llaman miedo".

El más fuerte lo acompañó al Salón del Trono. El monarca **expuso/exponía** (19) las condiciones que ya habían escuchado otros candidatos. Dijo:

— "Si tú consigues pasar tres noches seguidas en el castillo, derrotar a los espíritus y devolverme mi tesoro, **habrá/habrán** (20) dos semanas de fiestas en tu honor, te concedo la mano de mi amada y bella hija, y la mitad de mi reino como dote".

Juan replicó:

— "Se lo agradezco, Su Majestad, pero yo sólo vine para saber lo que es el miedo."

"Qué hombre tan valiente, qué honesto", pensó el rey, "pero ya guardo pocas esperanzas de recuperar mis dominios, ya **ha habido/han habido** (21) tantos que lo han intentado."

Juan sin Miedo se fue al castillo y escogió uno de los 200 cuartos que **había/habían** (22) ahí. Colgó sus hachas de la pared, pensando "nunca se sabe, y así siempre voy a tenerlas cerca" y se acostó. A medianoche, lo despertó un alarido muy alto.

— "¡Uhhhhhhhhh! Un espectro se deslizaba sobre el suelo sin tocarlo".

— "¿Quién eres tú, que te **atreves/atrevas** (23) a despertarme?" Preguntó Juan.

Un nuevo alarido por respuesta, y Juan Sin Miedo le tapó la boca con una bandeja que adornaba la mesa. El espectro se quedó mudo y se desapareció en el aire.

A la mañana siguiente el rey visitó a Juan Sin Miedo y pensó:

"Es sólo una pequeña batalla. Aún quedan dos noches".

Pasó el día y se fue el sol. Como la noche anterior, Juan Sin Miedo se acostó, pero esta vez apareció un fantasma espantoso que **lanzó/lanzaba** (24) un bramido: ¡Uhhhhhhhhh! Juan Sin Miedo cogió una de sus hachas y cortó la cadena que el fantasma arrastraba. Al no estar sujeto, el fantasma se elevó y desapareció.

Al amanecer, el rey volvió a visitarlo y pensó:

"Nada de esto habrá servido si él no repite la hazaña una vez más".

Llegó el tercer atardecer, y después, la noche. Juan Sin Miedo ya **dormía/durmió** (25) cuando escuchó acercarse a una momia. Y preguntó:

—"Dime qué motivo tienes para interrumpir mi sueño."

Ya que no contestó, Juan agarró un extremo de la venda y tiró. Retiró todas las vendas y encontró a un mago, quien dijo:

— "No **hay/hain** (26) trucos de magia que valgan contra ti. Déjame libre y yo rompo el encantamiento".

Al amanecer, **había/habían** (27) muchas gentes en las puertas del castillo, y cuando apareció Juan Sin Miedo el rey dijo: "¡Voy a cumplir mi promesa y más! ¡No **va a haber/van a haber** (28) dos sino cuatro semanas de fiesta!" Pero acá no acabó la historia: Cierto día en que el ahora príncipe dormía, la princesa decidió sorprenderle regalándole una pecera. Pero **tropezó/tropezaba** (29) al inclinarse, y el contenido, agua y peces cayeron sobre la cama que ocupaba Juan.

—" ¡Ahhhhhh!" exclamó Juan al sentir los peces en su cara - ¡Qué miedo!

La princesa rompió a reír, ya que no **había/habían** (30) peligros, espectros o espantos que **asustaban/asustaran** (31) a Juan, pero él sí les cogió miedo a unos simples peces de colores. Le dijo, riendo todavía:

— "No tengas miedo, te voy a guardar el secreto."

Y así fue, y todavía se le conoce como Juan Sin Miedo.

Appendix D

Questionnaire-reading task

1. *María engañó a su novio. Una amiga común hace de intermediaria. Después de haber hablado con el novio, Juan, dice:*
   Lo siento María, pero Juan dice _____ no quiere verte nunca jamás.
   a) de que b) que

2. *Un periodista entrevista a un pintor que acaba de presentar una serie de cuadros preciosos que son completamente diferentes de los que solía vender antes. Además, resulta que algunos ya los hizo hace veinte años. Pregunta el periodista: ¿Por qué usted esperó tanto antes de presentarnos estas obras? Contesta el artista:*
   Porque pensaba, y todavía pienso, que en aquel momento no _____las críticas tan positivas que estas obras están recibiendo ahora.
   a) pudo haber b) pudieron haber

3. *Un abuelo está contándoles a sus nietos de su niñez. Uno de ellos, ansioso de saber de estos tiempos pasados, pregunta:*
   ¿Papi, cuando usted era niño, ¿acá ya _____ (1) carros?
   Contesta el abuelo:
   ¡Claro que los_____ (2), no soy tan viejo!
   (1) a) había b) habían
   (2) a) había b) habían

4. *A Inés le acaban de robar el carro, que tenía aparcado en algún callejón obscuro. Aunque no es la cosa más sensata que se pueda hacer, una amiga trata de consolarla diciendo:*
   No es culpa tuya, es que siempre_____ unas personas malas.
   a) habrá b) habrán

5. *Desde pequeño, Francisco ha soñado con mudarse a Madrid. Ahora, su empresa le anunció que, cuando él quiera, lo pueden transferir a la sucursal de esta ciudad. Le dice a su madre:*
   Dentro de dos años, _____a Madrid, ya compré una casa allí.
   a) voy a mudarme b) me mudo c) me mudaré

6. *Dos personas están en un evento en que se presentan los carros del año. Una de ellas pretende que son exactamente los mismos que el año pasado. La otra persona dice:*
   No, no, estás equivocado. Por ejemplo, el año pasado no_____los carros amarillos que vimos antes.
   a) hubo b) hubieron

7. *Ana tiene problemas amorosos bastante serios. Después de semanas de sentirse muy mal, le cuenta todo a su prima.*
   Yo me esforcé muchísimo e intenté ser perfecta para él pero no logré nada. Él estaba como confuso y aunque _____(1) veces que me trataba bien, también _____(2) muchas veces que estaba distante y antipático conmigo.
   (1) a) hubiera b) hubieran
   (2) b) había b) habían

8. *Un domingo, dos muchachos se encuentran. Pregunta el primero: ¿Tú y la familia van a la playa esta tarde? Contesta el otro:*
   _____Mi viejo está enfermo.
   a) No, hoy no vamos, no. b) No, hoy no vamos.

9. *Marlén está leyéndole a su hija una historia de horror sobre gallinas posesas que se comen niños. Por supuesto, la hija le coge miedo y dice: "Mami, tengo miedo". Contesta Marlén:*
   No te preocupes, acá nunca _____las gallinas de que habla el libro.
   a) ha habido b) han habido

10. *Unos amigos invitaron a Marilyn y Julio a cenar. Julio es una de estas personas que nunca quiere llegar con las manos vacías. Marilyn le había prometido cocinar un bizcocho para regalárselo a los amigos. Cuando están por irse, Julio le pregunta a Marilyn: ¿Y en dónde es que está el bizcocho? Contesta Marilyn:*
    Bueno, no es que me olvidara de hacerlo, pero ya que estaba lloviendo tanto, no quise salir y en casa no_____los huevos necesarios para cocinarlo.
    a) había b) habían

11. *Ana y María están planeando una excursión. Por fin se pusieron de acuerdo sobre la destinación. Ana, a quien le gusta visitar museos, exposiciones, etc., dice:*
    Pues, déjame buscar toda la información y _____¿Vale?
    a) te llamo p'atrás b) vuelvo a llamarte

12. *Juan, un español, trata de convencerle a Tony, un amigo puertorriqueño, de que en España todo es mejor, lo cual este último no puede creer. Juan acaba de mencionar una pila de problemas que existen en Puerto Rico. Replica Tony:*
Tienes razón, pero en España_____los mismos tipos de problemas.
a) habrá b) habrán

13. *Dos personas están hablando de que ha aumentado el nivel de la pelota. Una tercera, más crítica, dice:*
La gente que dice que el nivel era más bajo en el pasado, no se acuerdan de los buenos juegos que veían cuando eran jóvenes. Tal vez no _____(1) los talentos que _____(2) hoy en día, pero también _____(3) muchos peloteros muy buenos.
(1) a) hubiera b) hubieran
(2) a) hay b) hayn
(3) a) había b) habían

14. *Después de un caso severo de contaminación, comenta un experto ante las cámaras de la prensa:*
La semana pasada, en esta presa _____miles de peces, veinte patos y tres garzas. Ahora está lleno de basura y los animales se han ido, o, peor aún, están muertos.
a) hubo b) hubieron

15. *El carro de Fernanda viene fallando desde hace tiempo. No son averías gordas, pero el carro tiene ya más de quince años y la joven no está segura de que las reparaciones, que pueden salir caras, valgan la pena. Le pregunta a su hermano:*
Yo no sé cuál es la mejor opción: cambiarlo o hacer que me lo arreglen.
a) ¿Qué tú harías? b) ¿Qué harías tú? c) ¿Qué harías?

16. *Tony está viendo las noticias. Después de acabado el programa, le pregunta a la novia:*
Oye ¿tú sabías que cada fin de semana_____veinte accidentes fatales en nuestra ciudad?
a) hay b) hayn

17. *Después de que los vecinos volvieran de una visita al zoológico. Marlén les pregunta* ¿_____(1) nuevos animales en el zoológico?
La familia, entusiasta, contesta:
¡Sí, sí los _____(2)! Vimos dos nuevos grupos de monos araña y un tigre que acababa de llegar de la India.
(1) a) hubo b) hubieron
(2) a) hubo b) hubieron

18. *Juan está contándole a su madre que a la hermana, María, le explotó una goma en la carretera. Pregunta la madre: ¿Qué ella hizo entonces? Contesta Juan:*
Llamó al esposo_____a cambiarla.
a) para que él viniera b) para él venir c) para que viniera

19. *Dos personas están hablando de literatura. El primero tiene la impresión de que este año no salieron sino buenas novelas, lo que también es la opinión de la crítica literaria. El otro no está de acuerdo y dice:*
La gente que dice que este año no salieron sino buenos libros, no saben de qué hablan, porque siempre_____ (1) libros malos y libros buenos y siempre los_____ (2).
(1) a) ha habido b) han habido
(2) b) habrá b) habrán

20. *Juanito está llenando un crucigrama con la ayuda de su mamá. Después de un rato, la madre le dice al muchacho:*
Creo que ya _____(1) los suficientes indicios como para tú poder terminar el rompecabezas sin mi ayuda.
Contesta Juanito:
¡No mami, no los _____(2) todavía!
(1) a) hay b) hayn
(2) a) hay b) hayn

21. *Armando está hablando con su hijo, Juan, que nació a mediados de los años 80, sobre el día de su nacimiento. Dice:*
Recuerdo que, para entonces, _____las primeras víctimas de SIDA y estábamos como un poco preocupados, porque a tu mamá le tuvieron que poner sangre después del parto.
a) empezó a haber b) empezaron a haber

22. *María está hablando con su jefe, Julio, que acaba de pedirle que haga un trabajo importante el día siguiente. Sin embargo, María ya está metida en un proyecto que le toma mucho de su tiempo. Por ello, le contesta:*
Mañana_____este trabajo, pero puede ser que no me dé tiempo.
a) voy a hacer b) hago c) haré

23. *Dos niños fueron al parque con la abuela. Cuando vuelven a su casa, la mamá les pregunta: ¿Vieron palomas en el parque? Uno de los hermanitos contesta:*
Sí, sí _____como once.
a) debía haber b) debían haber

24. *La semana pasada, Ana estuvo muy enferma. Ahora ya se siente mejor y le cuenta a una amiga:* Ahora ya me siento un poquito mejor, pero la semana pasada, _____veces que yo tenía tanta fiebre que, estando en la propia cama, yo no sabía dónde carajo estaba. a) había b) habían
25. *Ana está en una guagua. Delante de ella, un hombre se levanta a bajarse. De pronto, Ana ve que hay un celular en el asiento. Se levanta y dice:* Señor, con permiso, _____ a) ¿este celular es suyo? b) ¿este es su celular?
26. *Fernanda está viendo las noticias. De golpe, le grita al esposo, que es chileno:* ¡Cariño, ven a ver esto, _____tres terremotos en Chile! a) acaba de haber b) acaban de haber
27. *Ana y Marilyn quieren organizar una cena. Pregunta Ana: ¿Quien más podríamos invitar? De golpe Marilyn se acuerda de dos chicas, María y María, a quienes conocieron una semana antes y dice:* Pues, _____esas dos Marías que conocimos el otro día. a) hay b) hayn
28. *Dos amigos están hablando del vegetarianismo. Uno de ellos dice que es de todos los tiempos. El otro replica:* Yo pienso que siempre_____(1) personas que respetan a los animales, pero no creo que siempre _____(2) vegetarianos. (1) a) ha habido b) han habido (2) a) haya habido b) hayan habido
29. *En la empresa donde trabaja Marilyn, hay un compañero nuevo. Los demás le están preguntando por sus experiencias laborales previas. Dice el nuevo:* _____ a esta empresa, estaba trabajando en Alemania. a) Antes de yo venir b) Antes de que yo viniera
30. *Dos hombres están hablando de literatura británica. Uno de ellos dice que él prefiere leer los libros en inglés. El otro, por el contrario, suele esperar hasta que se traduzcan al español. Dice el primero:* Pero entonces tú tienes que esperar muchísimo, ya que recuerdo haber leído que, por ejemplo, en Inglaterra sólo_____ los traductores necesarios para traducir un décimo de las novelas que se habían publicado en el 2008. a) había b) habían

31. *Alicia está describiéndole a su esposo cuánto ha cambiado su casa paterna desde su niñez. Dice:* No pienso que en aquel entonces ya _____las butacas que están en la sala, el armario que está en la habitación de mis papás y los cuadros que están en la pared del pasillo. a) hubiera b) hubieran
32. *Un maestro está hablándole a uno de sus amigos, Armando, de la importancia que, según él, tiene la educación. Dice:* Armando, los estudios son tan precisos como la comida, por ejemplo, si no se estudiara, no _____ los conocimientos de la anatomía humana que te salvaron la vida el año pasado. a) habría b) habrían
33. *Después de algún proyecto para mejorar la calidad del agua de las presas del país, un científico comenta:* Hace diez años, no_____ más de tres sapos en esta presa. Hoy en día, cuenta con veinte patos, tres garzas y miles de peces. a) hubo b) hubieron
34. *Los papás de Alicia organizan una fiesta, a la que van a asistir muchos amigos, de modo que necesitan de la ayuda de la joven. Dice el papá:* Lo siento Alicia, pero realmente te necesitamos aquí. Quiero que_____ con la fiesta. a) tú nos ayudes b) nos ayudes
35. *Armando, a quien le gusta mirar las estrellas, suele levantarse los domingos a las cuatro de la mañana para disfrutar de la vista que tiene en el balcón de su apartamento. Por la tarde, en la playa, le dice a su hermano:* Sobre las cuatro, ya _____carros en la calle. Qué raro, ¿verdad? a) empezó a haber b) empezaron a haber
36. *Después de que los vecinos volvieran de una visita al zoológico, Marlén les pregunta: ¿Qué animales vieron en el zoológico? Y ellos contestan, un poco desilusionados:* ¡Muy pocos!, ni siquiera_____los usuales grupos de leones, tigres y monos. a) había b) habían
37. *Iraida no encuentra las revistas que acaba de comprar. Le pregunta a Juan, su hermano, si él sabe dónde están. Contesta:* Ah sí, las dejé en la sala para que _____. a) se las lea b) se lean

38. *Ana tiene problemas amorosos bastante serios. Después de semanas de sentirse muy mal, le cuenta todo a su prima.*
¡Ay muchacha! Tienes que ayudarme. Estoy enamorada de mi mejor amigo o, mejor dicho, del que era mi mejor amigo, porque _____otros sentimientos, que están destruyendo la amistad.
a) empieza a haber b) empiezan a haber

39. *Al inicio del año escolar, una madre le pregunta al maestro*:
Maestro, con permiso ¿ _____ treinta alumnos en la clase de mi hijo, como el año pasado?
a) seguirá habiendo b) seguirán habiendo

40. *Marilyn, que es una vegetariana convencida. Dice:*
_____me tendrán que obligar a la fuerza.
a) Para yo comer carne, b) Para que yo coma carne,

41. *Juan, a quien invitaron a una fiesta a la que no pudo asistir, le pregunta al amigo que la organizó: ¿Qué tal la fiesta que organizaste? Éste contesta:*
¡Qué mal estaba! sólo sobre la una de la mañana _____más de dos invitados.
a) empezó a haber b) empezaron a haber

42. *Una familia está rumbo al zoológico. En el carro, dice la madre:*
¡Ojalá _____(1)esos leones que vimos la vez pasada!
Pregunta el esposo:
¿Por qué no los _____(2)?
(1) a) haya b) hayan
(2) a) habría b) habrían

43. *Dos amigos están hablando de excursiones. Uno de ellos dice:*
Nunca _____en La Parguera, pero me gustaría ir este verano.
a) estuve b) he estado

44. *Dos niñas fueron al parque con la abuela. Cuando vuelven a su casa, la mamá les pregunta: ¿Vieron ardillas en el parque? Una de las hermanas contesta:*
Sí, _____ como nueve.
a) habría b) habrían

45. *Armando, a quien le gusta mirar las estrellas, suele levantarse los domingos a las cuatro de la mañana para disfrutar de la vista que tiene en el balcón de su apartamento. Durante el desayuno, su esposa se queja de que no pudiera dormir por el ruido de la calle. Replica Armando:*
Qué raro, esta mañana no_____más carros que otros domingos.
a) había b) habían

References

Álvarez-Martínez, M. Á. (1996). Extremeño. In M. Alvar-López, *Manual de dialectología hispánica: El español de España* (pp. 171-182). Barcelona: Ariel.

Álvarez-Nazarío, M. (1991). *Historia de la lengua española en Puerto Rico: Su pasado y su presente en el marco de la realidad social.* San Juan, PR: Academia Puertorriqueña de la Lengua Española.

Abbott, B. (2004). Definiteness and indefiniteness. In L. Horn, & G. Ward, *The handbook of pragmatics* (pp. 122-149). Oxford: Blackwell.

Alba, O. (2000). *Nuevos aspectos del español en Santo Domingo.* Santo Domingo: La Trinitaria.

Alba, O. (2004). *¿Cómo hablamos los dominicanos? Un enfoque sociolingüístico.* Santo Domingo: Grupo León Jimenes.

Allen, K., Pereira, F., Botvinick, M., & Goldberg, A. E. (2012). Distinguishing grammatical constructions with fMRI pattern analysis. *Brain & Language*, *123*, 174-182.

Almeida, M., & Hernández-Campoy, J. M. (2005). *Metodología de la investigación sociolingüística.* Málaga: Comares.

Alvar-López, M. (2000). *El español de la República Dominicana: Estudios, encuestas, textos.* Alcalá de Henares: La Goleta.

Andersen, H. (1973). Abductive and deductive change. *Language*, *49* (4), 765-793.

Aponte-Alequín, H. (2014). *Desafíos del español caribeño: El debate sobre el modo y la microvariación modal.* San Juan, PR: Universidad de Puerto Rico PhD dissertation.

Ashby, W. J., & Bentivoglio, P. (1993). Preferred argument structure in French and Spanish. *Language Variation and Change*, *5* (1), 61-76.

Baayen, H. R. (2008). *Analyzing linguistic data: A practical introduction to statistics using R.* Cambridge, MA: Cambridge University Press.

Baayen, H. R. (2014). Multivariate statistics. In R. Podesva, & D. Sharma, *Research methods in linguistics* (pp. 337-372). Cambridge, MA: Cambridge University Press.

Bailey, G. (2002). Real and apparent time. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes, *The handbook of language variation and change* (pp. 312-332). Oxford: Blackwell.

Bello, A. (1860). *Gramática de la lengua castellana al uso de los americanos.* Bogotá: Echeverría Hermanos.

Bentivoglio, P., & Sedano, M. (1989). *Haber*: ¿Un verbo impersonal? Un estudio sobre el español de Caracas. In G. De Granda, *Estudios sobre el español de América y lingüística afroamericana: Ponencias presentadas en el 45 Congreso internacional de americanistas* (pp. 59-81). Bogotá: Instituto Caro y Cuervo.

Bentivoglio, P., & Sedano, M. (1996). Venezuela. In M. Alvar-López, *Manual de dialectología hispánica: El español de América* (pp. 116-133). Barcelona: Ariel.

Bentivoglio, P., & Sedano, M. (2011). Morphosyntactic variation in Spanish-speaking Latin America. In M. Díaz-Campos, *The handbook of Hispanic sociolinguistics* (pp. 123-147). Oxford: Blackwell.

Beutel, A. M., & Mooney-Marini, M. (1995). Gender and values. *American Sociological Review*, *60* (3), 436-448.

Birner, B. (1994). Information status and word order. An analysis of English inversion. *Language*, *70* (2), 233-259.

Birner, B., & Ward, G. (1996). A crosslinguistic study of postposing in discourse. *Language and Speech*, *39* (2-3), 113-142.

Blas-Arroyo, J. L. (1995-1996). A propósito de un caso de convergencia gramatical por causación múltiple en el área de influencia lingüística catalana: Análisis sociolingüístico. *Cuadernos de Investigación Filológica*, *21-22*, 175-200.

Blas-Arroyo, J. L. (1999). *Lenguas en contacto: Consecuencias lingüísticas del bilingüismo social de las comunidades de habla del este peninsular.* Frankfurt am Main/Madrid: Vervuert/Iberoamericana.

Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Activation or implicit learning? *Journal of Experimental Psychology: General*, *129* (2), 177-192.

Bock, K., Dell, G. S., Chang, F., & Onishi, K. (2007). Persistent structural priming from language comprehension to language production. *Cognition*, *104*, 437–458.

Bolinger, D. (1954). Further comment on *haber*. *Hispania*, *37* (3), 334-335.

Bolinger, D. (1977). *Meaning and form.* London/New York, NY: Longman.

Boyd-Bowman, P. (1976). Patterns of Spanish emigration to the Indies until 1600. *Hispanic American Historical Review*, *56* (4), 580-604.

Britain, D. (1992). Linguistic change in intonation: The use of high rising terminals in New Zealand English. *Language Variation and Change*, *4* (1), 77-104.

Brown, E., & Rivas, J. (2012). Grammatical relation probability: How usage patterns shape analogy. *Language Variation and Change*, *24* (3), 317-341.

Bucholtz, M. (2009). From stance to style: Gender, interaction, and indexicality in Mexican immigrant youth slang. In A. Jaffa, *Stance: Sociolinguistic perspectives* (pp. 146-170). Oxford: Oxford University Press.

Bull, W. (1943). Related functions of *haber* and *estar*. *The Modern Language Journal*, *27* (2), 119-123.

Bull, W., Cantón, A., Cord, W., Farley, R., Finan, J., Jacobs, S., Jaeger, R., Koons, M., & Teugel, B. (1947). Modern Spanish verb-form frequencies. *Hispania*, *30* (4), 451-466.

Bybee, J. (2001). *Phonology and language use.* Cambridge, MA: Cambridge University Press.

Bybee, J. (2003a). Cognitive processes in grammaticalization. In M. Tomasello, *The new psychology of language: Cognitive and functional approaches to language* (Vol. 2, pp. 145-167). London: Laurens Erlbaum Associates.

Bybee, J. (2003b). Mechanisms of change in grammaticization. In B. D. Joseph, & R. D. Janda, *The handbook of historical linguistics* (pp. 602-623). Oxford: Blackwell.

Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language*, *82* (4), 711-733.

Bybee, J. (2007). Diachronic linguistics. In D. Geeraerts, & H. Cuyckens, *The Oxford handbook of cognitive linguistics* (pp. 945-987). Oxford: Oxford University Press.

Bybee, J. (2009). Grammaticization: Implications for a theory of language. In J. Guo, E. Lieven, N. Budwig, S. Ervin-Tripp, S. Ozcaliskan, & K. Nakamura, *Cross-linguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin* (pp. 345-355). New York, NY: Psychology Press.

Bybee, J., & Beckner, C. (2010). Usage-based theory. In B. Heine, & H. Narrog, *The Oxford handbook of linguistic analysis* (pp. 827-856). Oxford: Oxford University Press.

Bybee, J., & Pagliuca, W. (1985). Cross-linguistic comparison and the development of grammatical meaning. In J. Fisiak, *Historical semantics: Historical word-formation* (pp. 59-83). Berlin/New York, NY: Mouton De Gruyter.

Caravedo, R. (1996). Perú. In M. Alvar-López, *Manual de dialectología hispánica: El español de América* (pp. 152-168). Barcelona: Ariel.

Castillo-Trelles, C. (2007). La pluralización del verbo *haber* impersonal en el español yucateco. In J. Holmquist, A. Lorenzino, & L. Sayahi, *Selected proceedings of the third workshop on Spanish sociolinguistics* (pp. 74-84). Somerville, MA: Cascadilla.

Catalán, D. (1989). *El español: Orígenes de su diversidad.* Madrid: Paraninfo.

Cedergren, H., & Sankoff, D. (1974). Variable Rules: Performance as a statistical reflection of competence. *Language*, *50* (2), 333-355.

Chafe, W. (1974). Language and consciousness. *Language*, *50* (1), 111-133.

Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N. Li, *Subject and topic* (pp. 27-55). New York, NY: Academic Press.

Chambers, J. K. (2002). An informal epistemology. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes, *The handbook of language variation and change* (pp. 3-14). Oxford: Blackwell.

Chambers, J. K. (2009). *Sociolinguistic theory: Linguistic variation and its social significance.* Oxford: Blackwell.

Cheshire, J. (2002). Sex and gender in variationist research. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes, *The handbook of language variation and change* (pp. 423-443). Oxford: Blackwell.

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge, MA: MIT University Press.

Chomsky, N. (1995). *The Minimalist Program.* Cambridge, MA: The MIT Press.

Cienki, A. (2007). Frames, idealized cognitive models, and domains. In D. Geeraerts, & H. Cuyckens, *The Oxford handbook of cognitive linguistics* (pp. 170-187). Oxford: Oxford University Press.

Claes, J. (2011). ¿Constituyen las Antillas y el Caribe continental una sola zona dialectal? Datos de la variable expresión del sujeto pronominal en San Juan de Puerto Rico y Barranquilla, Colombia. *Spanish in Context*, *8* (2), 191-212.

Clark, E. V. (1978). Locationals: Existential, locative, and possessive constructions. In J. H. Greenberg, C. H. Ferguson, & E. A. Moravcsik, *Universals of human language. Volume 4: Syntax* (pp. 85-126). Stanford, CA: Stanford University Press.

Clark, L. (2007). Cognitive sociolinguistics: A viable approach to variation in linguistic theory. In W. J. Sullivan, & A. R. Lommel, *Lacus forum 33: Variation* (pp. 1-14). Houston, TX: Lacus.

Clark, L. (2008). Re-examining vocalic variation in Scottish English: A Cognitive Grammar approach. *Language Variation and Change*, *20* (2), 255-273.

Company-Company, C. (2003). La gramaticalización en la historia del español. *Medievalia*, *35*, 3-61.

Comrie, B. (1989). *Language universals and linguistic typology.* Chicago, IL: Chicago University Press.

Corbella-Díaz, D. (1996). Fuentes del vocabulario canario: Los préstamos léxicos. In J. Medina-López, & D. Corbella-Díaz, *El español de Canarias hoy: Análisis y perspectivas* (pp. 105-142). Frankfurt am Main/Madrid: Vervuert/Iberoamericana.

Croft, W. (2000). *Explaining language change: An evolutionary perspective.* London/New York, NY: Longman.

Croft, W. (2003). *Typology and universals.* Cambridge, MA: Cambridge University Press.

Croft, W. (2007). Construction Grammar. In D. Geeraerts, & H. Cuyckens, *The Oxford handbook of cognitive linguistics* (pp. 463-508). Oxford: Oxford University Press.

Croft, W., & Cruse, A. (2004). *Cognitive linguistics.* Cambridge, MA: Cambridge University Press.

Cruz-Arroyo, M. D. (2000). *Mil años de historia de España.* Madrid: Alianza.

D'Aquino Ruiz, G. (2004). *Haber* impersonal en el habla de Caracas: Análisis sociolingüístico. *Boletín de Lingüística*, *21*, 3-26.

D'Aquino-Ruiz, G. (2008). El cambio lingüístico de *haber* impersonal. *Núcleo*, *20* (25), 103-124.

Davies, M. (2002-). *Corpus del Español. 100 million words (1200s-1900s)*. Retrieved October 2013, from http://www.corpusdelespanol.org/.

Davies, M. (2008-). *Corpus of Contemporary American English. 450 million words, 1990-2012*. Retrieved October 2013, from http://corpus.byu.edu/coca/.

De Granda, G. (1994). *Español de América, español de África y hablas criollas hispánicas: Cambios, contactos y contextos.* Madrid: Gredos.

Delbecque, N. (2002). A Construction Grammar approach to transitivity in Spanish. In K. Davidse, & B. Lamiroy, *The nominative and accusative and their counterparts* (pp. 81-130). Amsterdam/Philadelphia, PA: John Benjamins.

DeMello, G. (1991). Pluralización del verbo *haber* impersonal en el español hablado culto de once ciudades. *Thesaurus*, *46*, 445-471.

Díaz-Campos, M. (1999-2000). La pluralización del verbo *haber* en dos áreas dialectales de Hispanoamérica. *Anuario de Lingüística Hispánica*, *15-16*, 235-245.

Díaz-Campos, M. (2003). The pluralization of *haber* in Venezuelan Spanish: A sociolinguistic change in real time. *IU Working Papers in Linguistics*, *3*, Article 5.

Dixon, R. M. W. (1979). Ergativity. *Language, 55* (1), 59-138.

Domínguez, C. L., Guzmán, B., Moros, L., Pabón, M., & Vilaín, R. (1998). Personalización de *haber* en el español de Mérida. *Lengua y Habla*, *3* (1), 23-36.

Domínguez-Hernández, M. A. (2007). Palabras preliminares. In M. A. Domínguez-Hernández, *La lengua en Cuba: Estudios* (pp. 7-26). Santiago de Compostela: Universidade de Santiago de Compostela.

Du Bois, J. (1987). The discourse basis of ergativity. *Language*, *63* (4), 805-855.

Eckert, P. (1989). The whole woman: Sex and gender differences in variation. *Language Variation and Change*, *1* (3), 245-267.

Eckert, P. (2008). Variation and the indexical field. *Journal of Sociolinguistics*, *12* (4), 453-476.

Eckert, P. (2009). Where does the social stop? In F. Gregersen, J. K. Parrott, & P. Quist, *Language variation - European perspectives III. Selected papers from the fifth international conference on language variation in Europe (ICLaVE 5), Copenhagen, June 2009* (pp. 13-30). Amsterdam/Philadelphia, PA: John Benjamins.

Epstein, C. F. (2007). Great divides: The cultural, cognitive, and social bases of the global subordination of women. *American Sociological Review*, *72* (1), 1-22.

Fauconnier, G. (2007). Mental spaces. In D. Geeraerts, & H. Cuyckens, *The Oxford handbook of cognitive linguistics* (pp. 351-376). Oxford: Oxford University Press.

Fauconnier, G., & Turner, M. (1996). Blending as a central process of grammar. In A. E. Goldberg, *Conceptual structure, discourse, and language* (pp. 113-129). Stanford, CA: Center for the Study of Language and Information Publications.

Fauconnier, G., & Turner, M. (2006). Conceptual integration networks. In D. Geeraerts, *Cognitive linguistics: basic readings* (pp. 303-372). Berlin/New York, NY: Mouton De Gruyter.

Feagin, C. (2002). Entering the community: Fieldwork. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes, *The handbook of language variation and change* (pp. 20-39). Oxford: Blackwell.

Fernández, F. (1982). Actitudes lingüísticas: Un sondeo preliminar. In O. Alba, *El español del Caribe: Ponencias del VI simposio de dialectología* (pp. 89-104). Santiago de Los Caballeros: Universidad Católica Madre y Maestra.

Fernández-Soriano, O. (1999). Two types of impersonal sentences in Spanish: Locative and dative subjects. *Syntax*, *2* (2), 101-140.

Fernández-Soriano, O., & Táboas-Baylín, S. (1999). Construcciones impersonales no reflejas. In I. Bosque, & V. Demonte, *Gramática descriptiva de la lengua española. Volumen 2: Las construcciones sintácticas fundamentales* (pp. 1723-1778). Madrid: Espasa-Calpe.

Fillmore, C. (1989). Grammatical construction theory and the familiar dichotomies. In R. Dietrich, & C. Graumann, *Language processing in social context* (pp. 17-38). Amsterdam: Elsevier Science.

Fillmore, C. (2006). Frame semantics. In D. Geeraerts, *Cognitive linguistics: Basic readings* (pp. 373-400). Oxford: Oxford University Press.

Fillmore, C., Kay, P., & O'Connor, M. C. (1988). Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language*, *64* (3), 501-538.

Flórez, L. (1946). Reseña a "Charles E. Kany, *American-Spanish syntax*, Chicago, University of Chicago Press, 1945. XII-466 págs.". *Thesaurus*, *2* (2), 372-385.

Fontanella de Weinberg, M. B. (1987). *El español bonaerense.* Buenos Aires: Hachette.

Fontanella de Weinberg, M. B. (1992a). *El español de América.* Madrid: Mapfre.

Fontanella de Weinberg, M. B. (1992b). Variación sincrónica y diacrónica de las construcciones con *haber* en el español americano. *Boletín de Filología de la Universidad de Chile*, *33*, 35-46.

Frago-Gracia, J. A. (1999). *Historia del español de América.* Madrid: Gredos.

Freeze, R. (1992). Existentials and other locatives. *Language*, *68* (3), 553-595.

Freites-Barros, F. (2003). Actitudes lingüísticas en torno a la pluralización de *haber* impersonal en los Andes venezolanos. *Interlingüística*, *14*, 375-382.

Freites-Barros, F. (2004). Pluralización de *haber* impersonal en el Táchira: Actitudes lingüísticas. *Boletín de lingüística*, *22*, 32-51.

Freites-Barros, F. (2008). Más sobre la pluralización de *haber* impersonal en Venezuela: El estado Táchira. *Lingua Americana*, *12* (22), 36-57.

Fried, M., & Östman, J.-O. (2004). Construction Grammar: A thumbnail sketch. In M. Fried, & J.-O. Östman, *Construction Grammar in a cross-language perspective* (pp. 11-86). Amsterdam/Philadelphia, PA: John Benjamins.

Garachana-Camarero, M. (1997). Acerca de los condicionamientos cognitivos y lingüísticos de la sustitución de *aver* por *tener*. *Verba*, *24*, 203-235.

García-Yebra, V. (1983). ¿Complemento directo o sujeto con las formas unipersonales de *haber*? *Revista de Filología Española*, *63* (1-2), 33-72.

Geeraerts, D. (2006). A rough guide to cognitive linguistics. In D. Geeraerts, *Cognitive linguistics: Basic readings* (pp. 1-28). Oxford: Oxford University Press.

Gibbs, R. W. (2009). Why cognitive linguists should care more about empirical methods. In M. González-Márquez, I. Mittelberg, S. Coulson, & M. J. Spivey, *Methods in cognitive linguistics* (pp. 2-18). Amsterdam/Philadelphia, PA: John Benjamins.

Gili-Gaya, S. (1980). *Curso superior de sintaxis española.* Barcelona: Vox.

Givón, T. (1999). Generativity and variation: The notion 'rule of grammar' revisited. In B. MacWhinney, *The emergence of language* (pp. 81-114). New York, NY: Psychology Press.

Goldberg, A. E. (1995). *Constructions: A Construction Grammar approach to argument structure.* Chicago, IL: Chicago University Press.

Goldberg, A. E. (1996). Making one's way through the data. In M. Shibatini, & S. Thompson, *Grammatical constructions: Their form and meaning* (pp. 29-54). Oxford: Clarendon Press.

Goldberg, A. E. (1999). The emergence of the semantics of argument structure constructions. In B. MacWhinney, *The emergence of language* (pp. 197-212). London: Lawrence Erlbaum Associates.

Goldberg, A. E. (2001). Patient arguments of causative verbs can be omitted: The role of information structure in argument distribution. *Language Sciences, 23*, 503-524.

Goldberg, A. E. (2003). Constructions: A new theoretical approach to language. *Trends in Cognitive Science*, *7* (5), 219-224.

Goldberg, A. E. (2005a). Argument realization. In J.-O. Östman, & M. Fried, *Construction grammars: Cognitive grounding and theoretical extensions* (pp. 17-43). Amsterdam/Philadelphia, PA: John Benjamins.

Goldberg, A. E. (2005b). Constructions, lexical semantics, and the correspondence principle: Accounting for generalizations and subregularities in the realization of arguments. In N. Erteschik-Shir, & T. Rapoport, *The syntax of aspect: Deriving thematic and aspectual interpretation* (pp. 215-236). Oxford: Oxford University Press.

Goldberg, A. E. (2006a). *Constructions at work: The nature of generalization in language.* Oxford: Oxford University Press.

Goldberg, A. E. (2006b). The inherent semantics of argument structure: The case of the English ditransitive construction. In D. Geeraerts, *Cognitive linguistics: Basic readings* (pp. 401-437). Oxford: Oxford University Press.

Goldberg, A. E. (2009). The nature of generalization in language. *Cognitive Linguistics*, *20* (1), 93-127.

Goldberg, A. E. (2010). Verbs, constructions, and semantic frames. In M. Rappaport-Hovav, E. Doron, & I. Sichel, *Lexical semantics, syntax, and event structure* (pp. 39-51). Oxford: Oxford University Press.

Goldberg, A. E. (2011). Corpus evidence of the viability of statistical preemption. *Cognitive Linguistics*, *22* (1), 131–153.

Goldberg, A. E., & Ackerman, F. (2001). The pragmatics of obligatory adjuncts. *Language*, *77* (4), 798-814.

Goldberg, A. E., Casenhiser, D., & Sethuraman, N. (2004). Learning argument structure generalizations. *Cognitive Linguistics*, *15* (3), 289–316.

González-Calvo, J. M. (2002). Sintaxis y semántica: *Haber* impersonal en español. In S. Platero, & L. M. Velarde, *Pulchre, bene, recte: Homenaje al Prof. Fernando González-Ollé* (pp. 639-656). Pamplona: Universidad de Navarra.

González-Tapia, C. (1994). *Estudio sobre el habla culta dominicana (resultados parciales de una investigación): Con un estudio especial de los idiolectos de Bosch y Balaguer.* Santo Domingo: Universidad Autónoma de Santo Domingo.

Grondelaers, S., Speelman, D., & Geeraerts, D. (2009). A case for a cognitive corpus linguistics. In M. González-Márquez, I. Mittelberg, S. Coulson, & M. J. Spivey, *Methods in cognitive linguistics* (pp. 149-170). Amsterdam/Philadelphia, PA: John Benjamins.

Gómez-Torrego, L. (1994). *Impersonalidad gramatical: Descripción y norma.* Madrid: Arcos libros.

Gutiérrez-Grova, A. (2007). Para la historia de la lengua culta escrita: Aproximación a la sintaxis del *Papel periódico de la Havana*. In M. A. Domínguez-Hernández, *La lengua en Cuba: Estudios* (pp. 57-100). Santiago de Compostela: Universidade de Santiago de Compostela.

Hay, J., Warren, P., & Drager, K. (2006). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics*, *34*, 458-484.

Heine, B. (1997). *Possession: Cognitive sources, forces, and grammaticalization.* Cambridge, MA: Cambridge University Press.

Henríquez-Ureña, P. (1940/1982). *El español en Santo Domingo.* Santo Domingo: Taller.

Hernández-Alonso, C. (1996). Castilla la Vieja. In M. Alvar-López, *Manual de dialectología hispánica: El español de España* (pp. 197-212). Barcelona: Ariel.

Hernández-Díaz, A. (2006). Posesión y existencia: La competencia de *haber* y *tener* y *haber* existencial. In C. Company-Company, *Sintaxis histórica de la lengua española. Primera parte: La frase verbal* (Vol. 2, pp. 1055-1160). Mexico City: Universidad Nacional Autónoma de México/Fondo de Cultura Económica.

Herrero-Ruiz de Lozaiga, F. J. (2008). *Haber*, *ser* y *estar* como verbos de existencia en el siglo XV. In I. Olza-Moreno, M. Casado-Velarde, & R. González-Ruiz, *Actas del XXXVII simposio internacional de la sociedad española de lingüística.* Pamplona: Servicio de Publicaciones de la Universidad de Navarra. Retrieved from: http://www.unav.es/linguis/simposiosel/actas/.

Heylen, K., Tummers, J., & Geeraerts, D. (2008). Methodological issues in corpus-based cognitive linguistics. In G. Kristiansen, & R. Dirven, *Cognitive sociolinguistics: Language variation, cultural models, social systems* (pp. 91-128). Berlin/New York, NY: Mouton De Gruyter.

Hills, E. C., & Anderson, J. (1929). The frequency of the moods and tenses of verbs in recent Spanish plays. *Hispania*, *12* (6), 604-606.

Hollmann, W. B., & Siewierska, A. (2011). The status of frequency, schemas, and identity in cognitive sociolinguistics: A case study on definite article reduction. *Cognitive Linguistics*, *22* (1), 25-54.

Holmquist, J. (2008). Gender in context: Features and factors in men's and women's speech in rural Puerto Rico. In M. Westmoreland, & J. A. Thomas, *Selected proceedings of the fourth workshop on Spanish sociolinguistics* (pp. 17-35). Somerville, MA: Cascadilla Proceedings Project.

Hothorn, T., Hornik, K., Strobl, C., & Zeileis, A. (2014). *Party: A laboratory for recursive partytioning*. Retrieved from: http://cran.r-project.org/web/packages/party/index.html.

Hudson, R. (1996). *Sociolinguistics.* Cambridge, MA: Cambridge University Press.

Israel, M. (1996). The way constructions grow. In A. E. Goldberg, *Conceptual structure, discourse, and language* (pp. 217-230). Stanford, CA: Center for the Study of Language and Information Publications.

Jackendoff, R. (1997). Twistin' the night away. *Language*, *73* (3), 534-559.

Jackendoff, R. (2008). Construction after construction and its theoretical challenges. *Language*, *84* (1), 8-28.

Jiménez-Sabater, M. A. (1978). Estructuras morfosintácticas en el español dominicano: Algunas implicaciones sociolingüísticas. In H. López-Morales, *Corrientes actuales en la dialectología del Caribe: Actas de un simposio* (pp. 167-180). San Juan, PR: Universidad de Puerto Rico.

Jiménez-Sabater, M. A. (1984). *Más datos sobre el español de la República Dominicana.* Santo Domingo: Universidad Autónoma de Santo Domingo.

Johnson, D. E. (2009). Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass*, *3* (1), 359-383.

Johnson, D.E. (2014). *Rbrul version 2.22.* Retrieved from: http://www.danielezrajohnson.com/Rbrul.R.

Jorge-Morel, E. (1978). *Estudio lingüístico de Santo Domingo.* Santo Domingo: Taller.

Kany, C. E. (1945/1951). *Sintaxis hispanoamericana.* Madrid: Gredos.

Kay, P., & Fillmore, C. (1999). Grammatical constructions and linguistic generalizations: *The What's X doing Y?* construction. *Language*, *75* (1), 1-33.

Keenan, E. (1976). Towards a universal definition of subject. In C. N. Li, *Subject and topic* (pp. 305-333). New York, NY: Academic Press.

Keniston, H. (1937). *Spanish syntax list: A statistical study of grammatical usage in contemporary Spanish prose on the basis of range and frequency.* New York, NY: H. Holt and Company.

Kerswill, P. (1996). Children, adolescents, and language change. *Language Variation and Change*, *8* (2), 177-202.

Kerswill, P. (2002). Koineization and accommodation. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes, *The handbook of language variation and change* (pp. 669-702). Oxford: Blackwell.

Kerswill, P., & Williams, A. (2000). Creating a new koine: Children and language change in Milton Keynes. *Language in Society*, *29* (1), 65-115.

Kiesling, S. F. (2005). Variation, stance, and style: Word-final *-er*, high rising tone, and ethnicity in Australian English. *English World-Wide: A Journal of Varieties of English, 26* (1), 1-42.

Kiesling, S. F. (2009). Style as stance. In A. Jaffa, *Stance: Sociolinguistic perspectives* (pp. 171-194). Oxford: Oxford University Press.

Kristiansen, G. (2008). Style-shifting and shifting styles: A socio-cognitive approach to lectal variation. In G. Kristiansen, & R. Dirven, *Cognitive sociolinguistics: Language variation, cultural models, social systems* (pp. 45-90). Berlin/New York, NY: Mouton De Gruyter.

Labov, W. (1966/2006). *The social stratification of English in New York City.* Cambridge, MA: Cambridge University Press.

Labov, W. (1972). *Sociolinguistic patterns.* Philadelphia, PA: University of Pennsylvania Press.

Labov, W. (1982). Building on empirical foundations. In W. P. Lehmann, & Y. Malkiel, *Perspectives on historical linguistics* (pp. 17-92). Amsterdam/Philadelphia, PA: John Benjamins.

Labov, W. (1984). Field methods of the project on linguistic change and variation. In J. Baugh, & J. Sherzer, *Language in use* (pp. 28-53). Englewood Cliffs, NJ: Prentice Hall.

Labov, W. (1994). *Principles of linguistic change. Volume 1: Internal factors.* Oxford: Blackwell.

Labov, W. (2001). *Principles of linguistic change. Volume 2: Social factors.* Oxford: Blackwell.

Labov, W. (2010). *Principles of linguistic change. Volume 3: Cognitive and cultural factors.* Oxford: Wiley-Blackwell.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago, IL: Chicago University Press.

Langacker, R. W. (1987). *Foundations of Cognitive Grammar. Volume 1: Theoretical prequisites.* Stanford, CA: Stanford University Press.

Langacker, R. W. (1990). *Concept, image, symbol: The cognitive basis of grammar.* Berlin/New York, NY: Mouton De Gruyter.

Langacker, R. W. (1991). *Foundations of Cognitive Grammar. Volume 2: Descriptive application.* Stanford, CA: Stanford University Press.

Langacker, R. W. (1995). Possession and possessive constructions. In J. R. Taylor, & R. E. MacLaury, *Language and the cognitive construal of the world* (pp. 51-79). Berlin/New York, NY: Mouton De Gruyter.

Langacker, R. W. (1999). Losing control: Grammaticization, subjectification, and transparency. In A. Blank, & P. Koch, *Historical semantics and cognition* (pp. 147-175). Berlin/New York, NY: Mouton De Gruyter.

Langacker, R. W. (2007). Cognitive Grammar. In D. Geeraerts, & H. Cuyckens, *The Oxford handbook of cognitive linguistics* (pp. 421-462). Oxford: Oxford University Press.

Langacker, R. W. (2008). *Cognitive Grammar: A basic introduction.* Oxford: Oxford University Press.

Langacker, R. W. (2009). *Investigations in Cognitive Grammar.* Berlin/New York, NY: Mouton De Gruyter.

Langacker, R. W. (2010). Cognitive Grammar. In B. Heine, & H. Narrog, *The Oxford handbook of linguistic analysis* (pp. 87-110). Oxford: Oxford University Press.

Lapesa, R. (1981). *Historia de la lengua española.* Madrid: Gredos.

Lavandera, B. (1978). Where does the sociolinguistic variable stop? *Language in Society*, *7* (2), 171-182.

Lavandera, B. (1984). *Variación y significado.* Buenos Aires: Hachette.

Lipski, J. (1996). Contacto de criollos en el Caribe hispánico: Contribuciones al español bozal. *América Negra*, *11*, 31-60.

Lipski, J. (2004). The Spanish language of Equatorial Guinea. *Arizona Journal of Hispanic Cultural Studies*, *8*, 115-130.

Lipski, J. (2009). *A history of Afro-Hispanic language.* Cambridge, MA: Cambridge University Press.

Lope-Blanch, J. M. (1989). *Estudios de lingüística hispanoamericana.* Mexico City: Universidad Nacional Autónoma de México.

Lope-Blanch, J. M. (1996). México. In M. Alvar-López, *Manual de dialectología hispánica: El español de América* (pp. 81-89). Barcelona: Ariel.

López-Morales, H. (1980). Sobre la pretendida existencia y pervivencia del 'criollo' cubano. *Anuario de Letras*, *18*, 85-116.

López-Morales, H. (1983). *Estratificación social del español de San Juan de Puerto Rico.* Mexico City: Universidad Nacional Autónoma de México.

López-Morales, H. (1992). *El español del Caribe.* Madrid: Mapfre.

López-Morales, H. (1998). *La aventura del español en América.* Madrid: Espasa-Calpe.

Lüdtke, J. (1994). Diferenciación y nivelación del español en la expansión a Canarias y al Caribe en el período de orígenes. In J. Lüdtke, *El español de América en el siglo XVI: Actas del simposio del Instituto Ibero-Americano de Berlín, 23 y 24 de abril de 1992* (pp. 39-56). Frankfurt am Main/Madrid: Vervuert/Iberoamericana.

Luque-Moreno, J. (1978). En torno al sintagma <*haber* impersonal + sustantivo> y sus orígenes latinos. *Revista Española de Lingüística*, *8*, 125-127.

Lyons, J. (1967). A note on possessive, existential, and locative sentences. *Foundations of Language*, *3*, 390-396.

Mack, S. (2011). A sociophonetic analysis of /-s/ variation in Puerto Rican Spanish. In L. A. Ortiz-López, *Selected proceedings of the thirteenth Hispanic linguistics symposium* (pp. 81-93). Somerville, MA: Cascadilla.

Malaver, I. (1999). Estudio de la conciencia lingüística sobre *hubieron*. *Lingua Americana*, *3* (5), 26-42.

Maldonado de Guevara, A. L. (1980). Consideraciones sobre el español actual. *Anuario de Letras*, *18*, 5-61.

Martín-Butragueño, P. (1999). ¿Es funcional la variación lingüística? In M. J. Serrano, *Estudios de variación sintáctica* (pp. 221-235). Frankfurt am Main/Madrid: Vervuert/Iberoamericana.

Medina-López, J. (1999). *El español de Canarias en su dimensión atlántica: Aspectos históricos y lingüísticos.* Valencia: Tirant Lo Blanch.

Megenney, W. W. (1999). El español afrocaribeño: ¿Mito o realidad? In M. Álvarez-Nazarío, & L. A. Ortiz-López, *El Caribe hispánico: Perspectivas lingüísticas actuales. Homenaje a Manuel Álvarez-Nazarío* (pp. 271-294). Frankfurt am Main/Madrid: Vervuert/Iberoamericana.

Meulleman, M., & Roegiest, E. (2012). Los locativos en la valencia de la construcción existencial española: ¿Actante o circunstante? *Zeitschrift für Romanische Philologie*, *128*, 57-70.

Michaelis, L., & Lambrecht, K. (1996). Toward a construction-based theory of language function: The case of nominal extraposition. *Language*, *72* (2), 215-247.

Milroy, L., & Gordon, M. (2003). *Sociolinguistics: Method and analysis.* Oxford: Blackwell.

Montes de Oca, M. (1994). La concordancia con *haber* impersonal. *Anuario de Letras*, *32*, 7-35.

Montes-Giraldo, J. J. (1982). Sobre el sintagma <*haber* + sustantivo>. *Thesaurus*, *32*, 383-385.

Morales, A. (1982). La perspectiva dinámica oracional en el español de Puerto Rico. In O. Alba, *El español del Caribe: Ponencias del VI simposio de dialectología* (pp. 205-219). Santiago de los Caballeros: Universidad Católica Madre y Maestra.

Morales, A. (1989). Hacia un universal sintáctico del español del Caribe: El orden SVO. *Anuario de Lingüística Hispánica*, *5*, 139-152.

Morales, A. (1997). La hipótesis funcional y la aparición de sujeto no nominal: El español de Puerto Rico. *Hispania*, *80* (1), 153-165.

Morales, A. (1999). Anteposición de sujeto en el español del Caribe. In M. Álvarez-Nazarío, & L. A. Ortiz-López, *El Caribe hispánico: Perspectivas lingüísticas actuales. Homenaje a Manuel Álvarez-Nazarío* (pp. 77-98). Frankfurt am Main/Madrid: Vervuert/Iberoamericana.

Moreno-Bernal, J. (1978). El uso de *haber* impersonal en un texto del siglo XIII. *Boletín de la Real Academia Española, 58* (214), 281-291.

Moreno de Alba, J. G. (1995). *El español en América.* Mexico City: Fondo de Cultura Económica.

Moreno-Fernández, F. (2003). *Metodología del Proyecto para el Estudio Sociolingüístico del Español de España y América (PRESEEA).* Retrieved from: http://www.linguas.net/portalpreseea.

Moya-Pons, F. (2008). *Historia del Caribe: Azúcar y plantaciones en el mundo atlántico.* Santo Domingo: Ediciones Ferilibro.

Navarro-Correa, M. (1992). Valoración social de algunas formas verbales en el habla de Valencia. *Lingüística Española Actual*, *14* (1), 97-106.

Navarro-Tomás, T. (1948). *El español en Puerto Rico: Contribución a la geografía lingüística hispanoamericana.* Río Piedras, PR: Universidad de Puerto Rico.

Norberg, D. (1968). *Manuel pratique de Latin médiéval.* Paris: Éditions A. & J. Picard et compagnie.

Nuño-Álvarez, M. (1996). Cantabria. In M. Alvar-López, *Manual de dialectología hispánica: El español de España* (pp. 183-196). Barcelona: Ariel.

Ortiz-López, L. A. (2000). El español de Puerto Rico en el contexto caribeño: Debate sobre su génesis. *Revista de Estudios Hispánicos*, *27* (1), 361-374.

Padrón, A. (1949). Giros sintácticos usados en Cuba. *Thesaurus*, *5* (1-3), 163-175.

Paolillo, J. C. (2013). Individual effects in variation analysis: Model, software, and research design. *Language Variation and Change*, *25* (1), 89-118.

Penny, R. (2006). *Gramática histórica del español.* Barcelona: Ariel.

Pereira-Scherre, M., & Naro, A. J. (1992). The serial effect on internal and external variables. *Language Variation and Change*, *4* (1), 1-13.

Pérez-Martín, A. M. (2007). Pluralización de *había* en el habla de El Hierro: Datos cuantitativos. *Revista de Filología de la Universidad de La Laguna*, *25*, 505-513.

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language, 39*, 633–651.

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin*, *134* (3), 427-459.

Picó, F. (2000). *Historia general de Puerto Rico.* San Juan, PR: Ediciones Huracán.

Poplack, S. (1984). Concord and sentential plural marking in Puerto Rican Spanish. *Hispanic Review, 52* (2), 205-222.

Preston, D. R. (2002). Language with an attitude. In J. K. Chambers, N. Schilling-Estes, & P. Trudgill, *The handbook of language variation and change* (pp. 40-66). Oxford: Blackwell.

Prince, E. (1992). The ZPG letter: Subjects, definiteness, and information status. In W. C. Mann, & S. A. Thompson, *Discourse description: Diverse linguistic analyses of a fund-raising text* (pp. 295-326). Amsterdam/New York, NY: John Benjamins.

Pütz, M., Robinson, J. A., & Reif, M. (2012). The emergence of cognitive sociolinguistics. *Review of Cognitive Linguistics*, *10* (2), 241-26.

Quilis, A. (1983). *La concordancia gramatical en la lengua española hablada en Madrid.* Madrid: Consejo Superior de Investigaciones Científicas.

Quintanilla-Aguilar, J. R. (2009). *La (des)pluralización del verbo* haber *existencial en el español salvadoreño: ¿Un cambio en progreso?* Miami, FL: University of Florida PhD dissertation.

Ramos-Alfajarín, J. R. (1998). Presentational sentences in Catalan dialects. *Catalan Working Papers in Linguistics*, *6*, 41-58.

Ramos-Alfajarín, J. R. (2001). El verb *haver-hi*: Evolució dels usos sintàctics. *Estudis Romànics*, *23*, 123-146.

R Core Team. (2013). *R: A language and environment for statistical computing.* Vienna: R Foundation for Statistical Computing. Retrieved from: http://www.r-project.org/.

Real Academia Española. (2005). *Diccionario panhispánico de dudas.* Madrid: Espasa-Calpe.

Real Academia Española. (2008a-). *Corpus diacrónico del español.* Retrieved June 2010, from: http://corpus.rae.es/cordenet.html.

Real Academia Española. (2008b-). *Corpus de referencia del español actual*. Retrieved October 2013, from: http://corpus.rae.es/creanet.html.

Real Academia Española, & Asociación de Academias de la Lengua Española. (2009). *Nueva gramática de la lengua española.* Madrid: Espasa-Calpe.

Rigau, G. (1993). El comportamiento sintáctico de los predicados existenciales en catalán. *Revista de Lenguas y Literaturas catalana, gallega y vasca*, *3*, 33-53.

Rivas, J., & Brown, E. (2012). Stage-level and individual-level distinction in morphological variation: An example with variable *haber* agreement. *Borealis: An International Journal of Hispanic Linguistics*, *1* (2), 73-90.

Rivas, J., & Brown, E. (2013). Concordancia variable con *haber* en español puertorriqueño. *Boletín de Lingüística*, *24* (37-38), 102-118.

Rodríguez-Mondoñedo, M. (2006). Spanish existentials and other accusative constructions. In C. Boeckx, *Minimalist essays* (pp. 326–394). Amsterdam/Philadelphia, PA: John Benjamins.

Sankoff, D., & Labov, W. (1979). On the uses of variable rules. *Language in Society*, *8*, 189-222.

Sankoff, G. (2004). Adolescents, young adults, and the critical period: Two case studies from *Seven Up*. In R. K. Macaulay, & C. Fought, *Sociolinguistic variation: Critical reflections* (pp. 121-139). Oxford: Oxford University Press.

Serrano, J. (2006). En torno a las vocales caedizas del español mexicano: Una aproximación variacionista. In P. Martín-Butragueño, *Líderes lingüísticos: Estudios de variación y cambio* (pp. 37-59). Mexico City: El Colegio de México.

Siegel, J. (1985). Koines and koineization. *Language in Society*, *14* (3), 357-378.

Silva-Corvalán, C. (2001). *Sociolingüística y pragmática del español.* Georgetown, D.C.: Georgetown University Press.

Silverstein, M. (2003). Indexical order and the dialectics of sociolinguistic life. *Language & Communication*, *23*, 193–229.

Stengaard, B. (2013). Subjects and objects with Latin *habere* and some of its Romance descendants. In I. Söhrman, & F. Josephson, *Diachronic and typological perspectives on verbs* (pp. 211-222). Amsterdam/Philadelphia, PA: John Benjamins.

Suñer, M. (1982). *Syntax and semantics of Spanish presentational sentences.* Georgetown, D.C.: Georgetown University Press.

Tagliamonte, S. (2002). Comparative sociolinguistics. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes, *The handbook of language variation and change* (pp. 729-763). Oxford: Blackwell.

Tagliamonte, S. (2006). *Analysing sociolinguistic variation.* Cambridge, MA: Cambridge University Press.

Tagliamonte, S. (2012). *Variationist sociolinguistics: Change, observation, interpretation.* Oxford: Wiley-Blackwell.

Tagliamonte, S., & Baayen, H. R. (2012). Models, forests, and trees of York English: *Was/were* variation as a case study for statistical practice. *Language Variation and Change*, *24* (2), 135-178.

Terrell, T. (1979). Final /s/ in Cuban Spanish. *Hispania, 62* (4), 599-612.

Terrell, T. (1982). Relexificación en el español dominicano: Implicaciones para la educación. In O. Alba, *El español del Caribe: Ponencias del VI simposio de dialectología* (pp. 301-318). Santiago de Los Caballeros: Universidad Católica Madre y Maestra.

Tomasello, M. (1995). Language is not an instinct. *Cognitive Development*, *10*, 131-156.

Tomasello, M. (2005). Beyond formalities: The case of language acquisition. *The Linguistic Review*, *22*, 183–197.

Tomasello. (2006). First steps toward a usage-based theory of language acquisition. In D. Geeraerts, *Cognitive linguistics: Basic readings* (pp. 439-458). Oxford: Oxford University Press.

Tomasello, M. (2007). Cognitive linguistics and first language acquisition. In D. Geeraerts, & H. Cuyckens, *The Oxford handbook of cognitive linguistics* (pp. 1092-1112). Oxford: Oxford University Press.

Torrego-Salcedo, E. (1999). El complemento directo preposicional. In I. Bosque, & V. Demonte, *Gramática descriptiva de la lengua española. Volumen 2: Las construcciones sintácticas fundamentales* (pp. 1779-1806). Madrid: Espasa-Calpe.

Travis, C. (2005). The yo-yo effect. Priming in subject expression in Colombian Spanish. In R. S. Gess, & E. J. Rubin, *Theoretical and experimental approaches to Romance linguistics: Selected papers from the third linguistic symposium on Romance languages, Salt Lake City, March 2004* (pp. 329-349). Amsterdam/Philadelphia, PA: John Benjamins.

Travis, C. (2007). Genre effects on subject expression in Spanish: Priming in narrative and conversation. *Language Variation and Change*, *19* (2), 101-135.

Trudgill, P. (1983). *On dialect.* Oxford: Blackwell.

Trudgill, P. (1986). *Dialects in contact.* Oxford: Blackwell.

Utley, J. (1954). *Haber* and *estar*. *Hispania*, *37* (2), 225.

Väänänen, V. (1967). *Introduction au Latin vulgaire.* Strasbourg: Centre de Philologie Romane, Université de Strasbourg.

Valdés-Bernal, S. (1994). El español de Cuba como parte del español del Caribe. In J. Lüdtke, & M. Perl, *Lengua y cultura en el Caribe hispánico: Actas de una sección del congreso de la asociación de hispanistas alemanes celebrado en Augsburgo, 4-7 de marzo de 1993* (pp. 1-13). Tübingen: Niemeyer.

Valdés-Bernal, S. (2007a). El poblamiento de Cuba y las áreas geolectales. In S. Valdés-Bernal, *Visión geolectal de Cuba* (pp. 7-28). Frankfurt am Main: Peter Lang.

Valdés-Bernal, S. (2007b). Las bases lingüísticas del español en Cuba. In M. A. Domínguez-Hernández, *La lengua en Cuba: Estudios* (pp. 27-56). Santiago de Compostela: Universidade de Santiago de Compostela.

Vaquero, M. (1978). Enseñar español, pero ¿qué español? *Boletín de la Academia Puertorriqueña de la Lengua Española*, *6*, 127-146.

Vaquero, M. (1996). Antillas. In M. Alvar-López, *Manual de dialectología hispánica: El español de América* (pp. 51-67). Barcelona: Ariel.

Waltereit, R., & Detges, U. (2008). Syntactic change from within and from without syntax: A usage-based analysis. In U. Detges, & R. Waltereit, *The paradox of grammatical change: Perspectives from Romance* (pp. 13 – 30). Amsterdam/Philadelphia, PA: John Benjamins.

Ward, G., & Birner, B. (1995). Definiteness and the English existential. *Language*, *71* (4), 722-742.

Weber, C. (1985). Three notes on *habeo* and *Ac* in the *Itinerarium Egeriae*. *Illinois Classical Studies, 10*, 285-294.

Weiner, J., & Labov, W. (1983). Constraints on the agentless passive. *Journal of Linguistics*, *19*, 29-58.

Weinreich, U., Labov, W., & Herzog, M. (1968). Empirical foundations for a theory of language change. In W. Lehman, & Y. Malkiel, *Directions for historical linguistics: A symposium* (pp. 97-195). Austin, TX: University of Texas Press.

Wolfram, W. (1986). Good data in a bad situation: Eliciting vernacular structures. In J. A. Fishman, A. Tabouret-Keller, M. Clyne, B. Krishnamurti, & M. Abdulaziz, *The Fergusionan impact: In honor of Charles A. Ferguson on the occasion of his 65th birthday. Volume 2: Sociolinguistics and the sociology of language* (pp. 3-22). Berlin/New York, NY: Mouton De Gruyter.

Wolfram, W., & Schilling-Estes, N. (2003). Dialectology and linguistic diffusion. In B. D. Joseph, & R. D. Janda, *The handbook of historical linguistics* (pp. 713-735). Oxford: Blackwell.

Index


Universiteit
Antwerpen